SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(95th Edition)

NOVEMBER 2020

BY: CENTRAL LIBRARY

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

PREFACE

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

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

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

This issue contains the information published during **December, 2020**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A Correlative Study of Centrality Measures across Real-World Networks, ***7.Harshita Rastogi*** and ***3.Minni Jain***, CSE, DTU

2. A Deep Language-independent Network to analyze the impact of COVID-19 on the World via Sentiment Analysis, ***6.Ashima Yadav*** and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

3. A Study on Combating Emerging Threat of Deepfake Weaponization, ***3.Rahul Katarya*** and ***8.Anushka Lal***, CSE, DTU

4. A study on Rocker-Bogie Suspension for a Planetary Rover Prototype, ***6.Anmol Singh*** and ***3.P K Jain***, Mechanical, DTU

5. A survey on enhanced RPL: Addressing the mobility in RPL, ***3.Rajesh Kumar Yadav*** and ***7.Nimisha Awasthi***, CSE, DTU

6. A theoretical comparison between CH3NH3PbI3 and CH3NH3SnI3 based solar cells, Gagandeep, ***3.Mukhtiyar Singh***, Ramesh Kumar, Vinamrita Singh and Sunita Srivastava, Physics, DTU

7. Accelerating Business Growth with Big Data and Artificial Intelligence, ***8.Awishkar Ghimire***[1], ***8.Surendrabikram Thapa***[1], ***8.Avinash Kumar Jha***[2], ***8.Surabhi Adhikari***[1] and ***8.Ankit Kumar***[2], [1]Computer and [2]Civil, DTU

8. An adaptive difference scheme for parabolic delay differential equation with discontinuous coefficients and interior layers, ***3.Aditya Kaushik*** and ***6.Nitika Sharma***, Applied Mathematics, DTU

9. Bot Detection in Social Networks Based on Multilayered Deep Learning Approach, Sandeep Singh Sengar, ***3.Sanjay Kumar***, ***8.Pradyot Raina*** and ***8.Mukul Mahaliyan***, CSE, DTU

10. Community detection in complex networks using network embedding and gravitational search algorithm, ***3.Sanjay Kumar***, B S Panda and ***8.Deepanshu Aggarwal***, CSE, DTU

11. Comparative Analysis of FPGA Based Hardware Design of Dynamic Time Warping Algorithm using Different Multiplier Architectures, ***6.Ritik Koul***, ***6.Mukul Yadav*** and ***3.Kriti Suneja***, Electronics, DTU

12. Control and Operation of a Small Autonomous DC Microgrid with Improved Dynamic Performance, Utsav Patel, ***3.Avirup Maulik*** and Debapriya Das, Electrical, DTU

13. Convenience store locations prioritization: a fuzzy TOPSIS-GRA hybrid approach, ***8.Jyotdeep Singh***, ***6.Parnika Tyagi***, ***3.Girish Kumar*** and ***3.Saurabh Agrawal***, Mechanical and Electrical, DTU

14. CVD Synthesis and Characterization of Ultrathin MoS2 Film, Mohammad Imran, Nahid Chaudhary, Aurangzeb Khurram Hafiz, ***3.Bharti Singh*** and Manika Khanuja, Physics, DTU

15. Depression and Impaired Mental Health Analysis from Social Media Platforms using Predictive Modelling T echniques, Vaibhav Jain, ***8.Dhruv Chandel***, Piyush Garg and ***3.Dinesh Kumar Vishwakarma***, IT, DTU

16. Design and Analysis of Rear Drive Axle of Go-Kart, ***8.Shubham Singh***, Mechanical, DTU

17. Design and Control of Medium-Voltage Multilevel Converter for Direct Grid Integration of Photovoltaic System, ***6.Jyoti Kulkarni***, ***3.Narendra Kumar*** and Bhim Singh Electrical, DTU

18. Dynamic Voltage Frequency Scaling in Multi-core Systems using Adaptive Regression Model, Manjari Gupta, Lava Bhargava and ***3.Indu Sreedevi***, Electronics, DTU

19. Effect of temperature on wear rate and coefficient of friction of Railway Brake Block, ***8.Vishal Kumar***, ***8.Nishant Yadav***, ***8.Prem Kumar*** and ***3.Roop Lal***, Mechanical, DTU

20. Effect of temperature on food consumption, immune system, antioxidant enzymes, and heat shock protein 70 of Channa punctata (Bloch, 1793), Samar Pal Singh, Tauqueer Ahmad, ***3.Jai Gopal Sharma*** and Rina Chakrabarti, Biotechnology, DTU

21. End to End Binarized Neural Networks for Text Classification, Kumar Shridhar, Harshil Jain, ***8.Akshat Agarwal*** and Denis Kleyko, Electrical, DTU

22. Enhanced piezoelectricity properties of reduced graphene oxide (RGO) loaded polyvinylidene fluoride (PVDF) nanocomposite films for nanogenerator application, Abhishek Anand, ***3.Deshraj Meena***, Kajal Kumar Dey and Mukesh Chander Bhatnagar, Physics, DTU

23. Estimation and Experimental Validation of Mean-Field Homogenised Effective Properties of Composite, ***6.A. Sharma***, ***3.V. Rastogi*** and ***3.A.K. Agrawal***, Mechanical, Production, Industrial and Automobile, DTU

24. Feature Selection Based Twin-Support Vector Machine for the Diagnosis of Parkinson's Disease, ***8.Surendrabikram Thapa***, ***8.Surabhi Adhikari***, ***8.Awishkar Ghimire*** and ***8.Anshuman Aditya***, Computer, DTU

25. Force and temperature analysis during distinct machining environment using an optimization approach, ***6.Anurag Sharma***, ***3.R C Singh***, ***3.Ranganath M Singari*** and S L Bhandarkar, Mechanical, DTU

26. Green Synthesis and Characterization of Copper Oxide Nanoparticles Using Psidium guajava Leaf Extract, ***7.Ammu Varughese***, ***3.Raminder Kaur*** and ***3.Poonam Singh***, Chemistry, DTU

27. Group operations and isomorphic relation with the 2-tuple linguistic variables, ***6.Tanya Malhotra*** and ***3.Anjana Gupta***, Applied Mathematics, DTU

28. Growth Story NDA versus UPA: Is Special Status of Indian States Justified?, ***3.K. V. Bhanu Murthy***, USME, DTU

29. Hardware Emulation of a Biorthogonal Wavelet Transform Based Heart Rate Monitoring Device, Ashish Kumar, Ramana Ranganatham, ***3.Manjeet Kumar*** and Rama Komaragiri, Electronics, DTU

30. Identification of the barriers of Reverse Supply Chain, ***8.Sunil***, ***8.Dharmanshu Kumar***, ***8.Udit Mehta***, ***3.Rakesh Kumar***, Mechanical, DTU

31. Impact of Supervised Classification Techniques for the Prediction of Student's Performance, ***8.Rahul*** and ***3.Rahul Katarya***, CSE, DTU

32. Laser Based Systems for Standoff Detection of CWA: A Short Review, M K Jindal, ***7.1.Mainuddin***, S Veerabuthiran and A K Razdan, Electronics, DTU

33. LIQUIDITY IN ASIAN FINANCIAL MARKETS: CROWDING OUT OR SPILLOVER EFFECT, ***3.SONAL THUKRALI*** and ***6.RAHUL SIKKA***, DSM, DTU

34. Melanoma Risk Prediction with respect to Modifiable Lifestyle Factors by Meta-Analysis Aided Machine Learning Technique, ***8.1.Avni Mehta***, ***8.1.Richa Nayak*** and ***3.Yasha Hasija***, Biotechnology, DTU

35. mTrust: Call Behavioral Trust Predictive Analytics Using Unsupervised Learning in Mobile Cloud Computing, ***7.1.Arka Bhowmik*** and Debashis De, CSE, DTU

36. Multi-Output Fast Wireless Charging for Electric Vehicle, ***8.Imroz Khan***, ***8.Anjali Kumari***, ***8.Hemant Sharma*** and ***3.Prof. Prem Prakash***, Electrical, DTU

37. Nanomaterials Based Biosensors for COVID-19 Detection-A Review, ***8.Sakshi Sharma***, ***8.Sonakshi Saini***, ***8.Maya Khangembam*** and ***3.Vinod Singh***, Physics, DTU

38. Numerical Simulation of n-TiO2/p-CIGS Solar Cell, Gagandeep, ***3.Mukhtiyar Singh***, Ramesh Kumar and Sonia Kumari, Physics, DTU

39. ON CONVEX DOMINANTS OF EXACT DIFFERENTIAL SUBORDINATION, ***3.S. SIVAPRASAD KUMAR*** AND ***6.SHAGUN BANGA***, APPLIED MATHEMATICS, DTU

40. On the performance of DF-based multi-hop system over α−κ−μand α−κ−μ-extreme fading channels, Tau Raphael Rasethuntsa, ***6.Manpreet Kaur***, SandeepKumar, Puspraj Singh Chauhan and Kuldeep Singh, CSE, DTU

41. Performance measurement of flexible manufacturing system in context to Bareilly camphor industry, ***6.Himani Saxena***, ***3.RS Mishra*** and ***3.AK Madan***, Mechanical Production Industrial and Automobiles, DTU

42. Performance evaluation of solar based combined pre-compression supercritical CO2 cycle and organic Rankine cycle, ***6.Yunis Khan*** and ***3.Radhey Shyam Mishra***, Mechanical, DTU

43. Phase evolution and microstructure of BaBi2Nb2O9 ferroelectric ceramics, ***6.Ankita Banwal*** and ***3.Renuka Bokolia***, Physics, DTU

44. Pre-Processing and Emoji Classification of WhatsApp Chats for Sentiment Analysis, ***8.Astha Mohta***, ***8.Atishay Jain***, ***8.1.Aditi Saluja*** and ***3.Sonika Dahiya***, CSE, DTU

45. Recognising Devanagari Script by Deep Structure Learning of Image Quadrants, ***3.Seba Susan*** and ***7.Jatin Malhotra***, IT, DTU

46. Review on Sheet Metal Formability, ***8.Shivang Chhikara***, ***8.Rajat Khatta***, ***8.Suraj Prakash Verma*** and ***8.Dadge Mukesh Shamrao***, Mechanical, DTU

47. RNA Secondary Structure Prediction using Machine Learning: A Review, ***7.Shubham Mittal*** and ***3.Yasha Hasija***, Biotechnology, DTU

48. Robust Stator Winding Fault Detection in PMSMs with Respect to Current Controller Bandwidth, Shaopo Huang, ***8.1.Anmol Aggarwal***, Elias G. Strangas, Kui Li, Feng Niu, and Xiaoyan Huang, Electrical, DTU

1. Vice Chancellor
1.1. *Ex Vice chancellor*
2. *Pro Vice Chancellor*
2.1. *Ex Pro Vice Chancellor*
3. *Faculty*
3.1. *Ex Faculty*
4. *Teaching-cum-Research Fellow*
4.1. *Asst. Librarian*
5. *Alumni*
6. *Research Scholar*
6.1. *Ex Research Scholar*
7. *PG Scholar*
7.1. *Ex PG Scholar*
8. *Undergraduate Student*
8.1. *Ex Undergraduate Student*

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

A Correlative Study of Centrality Measures across Real-World Networks

Harshita Rastogi
Department of Computer Science and Engineering
Delhi Technological University, DTU
Delhi, India
harshi.rastogi94@gmail.com

Minni Jain
Department of Computer Science and Engineering
Delhi Technological University, DTU
Delhi, India
minnijain@dtu.ac.in

***Abstract*—Centrality measures have evolved over the years and used over a variety of networks. Being one of the most basic measures to identify important nodes in ever-increasing modern-day networks, it has been manipulated and modified in every way possible to fit the requirement of the network and the way important nodes are perceived in it. Different centrality measures which seem to perform quite differently on a theoretical basis, provide similar results when applied to real-life networks. The centrality measures are studied based on the approach used, application areas, performance and measure the correlation among 14 centrality measures across 12 network topologies using Pearson, Spearman and Kendall correlation coefficients.**

Keywords—centrality measures, centrality measure correlation, real-world networks

I. INTRODUCTION

A network is a visual representation of similar data which hold some relation among them. It represents the topology of the system containing the data in the form of nodes as entities and edges as relations among them. This concept is used in various phases of real-world to depict the structure of any system in a more visually apprehensible form. Real-world networks can exist in a vast variety ranging from social, economic, biological networks etc. spread across various types e.g. bipartite, multi-layer, time-series etc. and various domains e.g. communication, bioinformatics, sciences etc. [1]. These networks are generally very large in size, thus making it difficult to analyze and interpret them. An efficient way to do so is by analyzing the important components of these networks. To effectively identify the important components, the concept of centrality measures is used. A node is regarded as central based on how much impact it lays on the network and how much the graph is affected after it is removed. Different types of networks have a different definition for node importance like structural significance, localization, connections and key users of the node etc [2]. In a more generalized perspective, centrality measures are highly influenced by the topology as well as its rules and payoff of the network [3].

A large variety of centrality measures have been introduced over the years to identify important nodes in different types of networks. This leaves the user with several options to choose and decide which measure(s) are most suitable for his approach in terms of efficiency, time and/or space complexity. Moreover, several centrality measures have shown to provide similar results despite their varying approaches. Few works in the field of finding the correlation among the centrality measures exists. Furthermore, the previously used works have either only compared basic centrality measures[4-6] or considered only a specific domain like sociometric networks [7] and biological networks [8,9] or have only used one measure to detect correlation [10]. In this paper, a comparison of centrality measures is being performed on two network datasets from 6 major real-world network categories applied over 14 different centrality measures. Although the centrality measures work on different principles, they show similar outcomes in case of real-work networks. This behavior causes redundancy and thus it is necessary to determine which centrality measures exhibit such behavior. 3 most commonly used methods have been applied namely Pearson, Spearman and Kendall correlation coefficients on each of the results to detect the degree of correlation among the centrality measures. This helps us to determine which centrality measures perform more efficiently and how correlated they are to each other.

This paper has been arranged as Section 2 contains the literature review of the real-world networks, centrality measures and correlation coefficients. Section 3 discusses the results obtained by the correlation and comparison between them. Finally, Section 4 presents the conclusion obtained from the correlation results.

II. METHOD USED

A set of 14 conceptually varying centrality measures have been analyzed across 12 medium-sized network datasets from 6 most common real-world network domains using 3 correlation coefficients to compare these centrality measures and determine the degree of correlation present between each pair of centrality measures. The three factors of our method are:

A. Real-world networks

These networks are generally decentralized or self-organized networks present in the real world. They can be man-made networks e.g. world wide web, powerline, airline, economic networks etc. or natural networks food web, protein, metabolic networks etc [11]. These networks possess specific structural properties like small “shortest path length”, high “clustering”,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

low cost and scale-free degree distribution [12]. The network topology largely affects the behavior of the centrality measures. To study this change in behavior, a set of two network structures have been used from each of biological, social, technological, temporal, road and power network domains. Table 1. describes these real-world networks in detail. All the network are openly available on the network repository[13]. The networks used are biological networks- Celegans network is a sparse network of substrates and the metabolic reactions among them and Diseasome network is a network representing people and the common set of diseases among them. Social Networks- Wiki-vote network is a Wikipedia network of who voted for whom till January 2008 and Facebook-ego network is a friendship network extracted from a Facebook social network. Technological networks- Firm-hi-tech network is a network of technical interactions among the users of a firm and Routers network is a set of routers and the communication paths among them. Temporal networks- Infect-hyper network is a human interaction network and Occupywallstnyc network is a twitter retweet network collected from various political and social hashtags. Road networks- Euro-road and Minnesota networks which are the road networks for a part of Europe and Minnesota respectively. Power networks- a 685-BUS network is a powerful network for the admittance matrix of a 685-Bus power system and BCSPWR-09 network is a US power network. The vertex and edge types describe the nature of the nodes and links between them. The vertices and edges are the numbers of nodes and links in the network. A maximum degree is the maximum number of links incident on a node in the network. Average degree is the average number of links incident on any node in the network. The density of the network is the number of connections the nodes hold out of the total possible number of connections. Average Clustering Coefficient is an average of the number of triangles presents locally. Assortivity is the degree of preference of a node to link to other nodes. The number of 3-cliques is the total number of triangles present in the graph. Maximum k-core is the maximal connected subgraph for the given network.

TABLE I. PROPERTIES OF NETWORK DATASETS

Networks	Vertex type	Edge type	Number of vertices	Number of edges	Density	Maximum degree	Average degree	Average clustering coefficient	Assortivity	Number of 3-cliques	Maximum k-core
Celegans	Substrates	Metabolic-reactions	453	2000	237	8	0.019779	0.646463	-0.225821	9900	11
Diseasome	People	Diseases	516	1200	50	4	0.008941	0.635830	0.066646	4100	11
Wiki-Vote	Voters, candidates	Votes	889	2900	102	6	0.007383	0.152847	-0.028776	6400	10
Facebook-ego	People	Friendships	2900	3000	769	2	0.000715	0.027247	-0.668214	273	4
Firm-hi-tech	Users	Interactions	33	124	28	9.19	0.235800	0.405000	-0.120000	454.5	8
Routers	Routers	Interactions	2100	6600	109	6	0.002972	0.246429	0.019161	31200	16
Infect-hyper	People	Proximities	113	2200	98	38	0.347029	0.534756	-0.122580	50600	29
Occupywall stnyc	Users	Retweets	3600	3900	3200	2	0.000604	0.037421	-0.691700	564	5
Euro-road	Locations	Roads	1200	1400	10	2	0.002057	0.167316	0.126684	96	3
Minnesota	Locations	Roads	2600	3300	5	2	0.000947	0.015960	-0.184813	159	3
685-BUS	Buildings	Routes	685	1300	12	3	0.005472	0.172478	0.181204	990	6
BCSPWR-09	Buildings	Routes	1700	2400	14	2	0.001614	0.075915	-0.100069	522	4

B. Centrality measures

There are more than 200 different types of centralities [14] proposed so far for different types of networks and based on different kinds of roles the nodes play in them. Due to limited resources and feasibility, the 14 major centrality measures have been studied. These centrality measures can be classified as *degree-based* e.g. degree, Hubble-index, lobby-index and diffusion centrality measures; *distance-based* e.g. Katz and laplacian centrality measures; *Eigen-based* e.g. alpha, eigenvector and clustering coefficient centrality measures; *neighborhood-based* e.g. PageRank and radiality centrality measures or *miscellaneous* e.g betweenness, information, local bridging centrality measures. Table 2. depicts a detailed description of the centrality measures including their formulae, application areas and approach used by them.

C. Correlation measures

Pearson, Spearman and Kendall correlation coefficients [15,16] have been used to measure correlations. All three methods measure the correlation between two variables on a scale between -1 to 1. On this scale, 0 means no correlation, -1 means negative correlation and 1 means positive correlation. The value of the correlation determines the strength of the relationship statistically. The description of the three methods is given below:

a) Pearson correlation coefficient:

It is the most commonly used method to find the strength and direction of the linear relationship between two continuous variables. It is a parametric correlation method as it depends on the data distribution. It uses mean and standard deviation to calculate the correlation. This method works on the assumptions that each observation holds a pair of values, each

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

variable is continuous and normally distributed, outliers should be absent, and the data is linear and homoscedas.

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 . \sum(Y - \bar{Y})^2}} \quad (1)$$

where $\bar{X}$ and $\bar{Y}$ are means of variables X and Y respectively.

b) Spearman correlation coefficient:

It is similar to the Pearson correlation coefficient with the difference that it uses ranks to sort the observations and calculate the similarity between them. It is a non-parametric rank-based method. Unlike the Pearson correlation, it works for both linear and non-linear observations. This makes it robust against outliers. It assumes the observation pairs to be independent and the relationship between the variables is monotonic. The variables can be measured on ordinal, interval or ratio scales

$$\rho = \frac{\sum_{i=1}^{n}(R(x_i) - \overline{R(x)})(R(y_i) - \overline{R(y)})}{\sqrt{\sum_{i=1}^{n}(R(x_i) - \overline{R(x)})^2 (R(y_i) - \overline{R(y)})^2}} \quad (2)$$

where $R(x_i)$ and $R(y_i)$ are the ranks of observations x_i and y_i respectively, $\overline{R(x)}$ and $\overline{R(y)}$ are the mean ranks of x and y respectively and n is the number of observation pairs.

c) Kendall correlation coefficient:

It is also a rank-based non-parametric correlation method similar to Spearman correlation coefficient. It can also be used for continuous and ordinal data. It uses the static dependence between the variables. It is used to represent the concordance and discordance number of pairs of the observations. It works on similar assumptions as Spearman correlation.

$$\tau = \frac{n_c - n_d}{n_c + n_d} = \frac{n_c - n_d}{n(n-1)/2} \quad (3)$$

where n_c, n_d are the concordant and discordant pairs respectively and n is the total number of pairs.

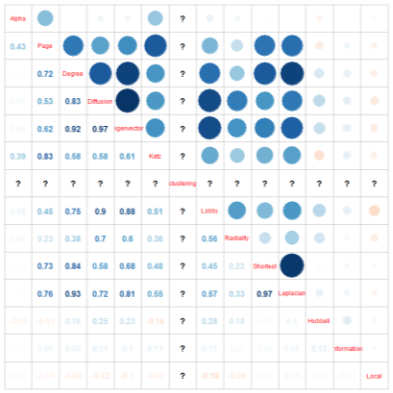
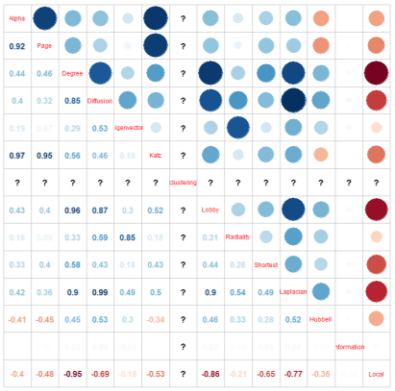
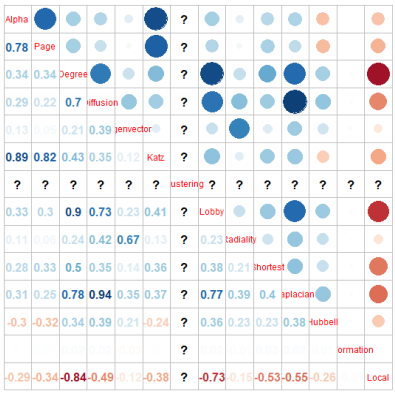
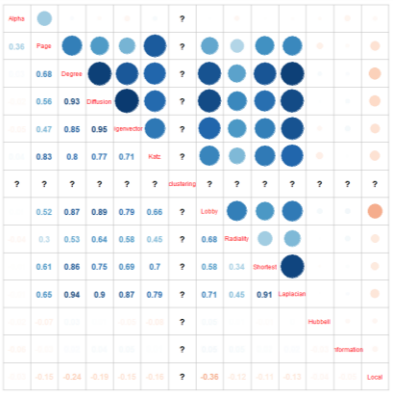
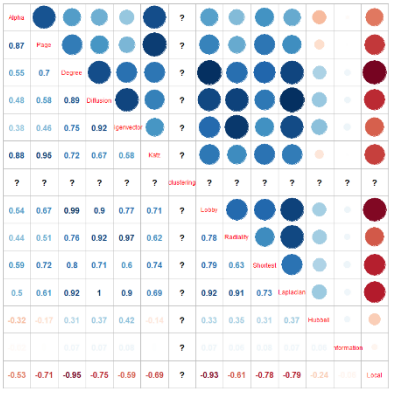
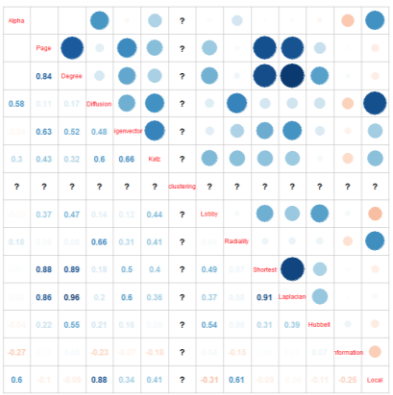
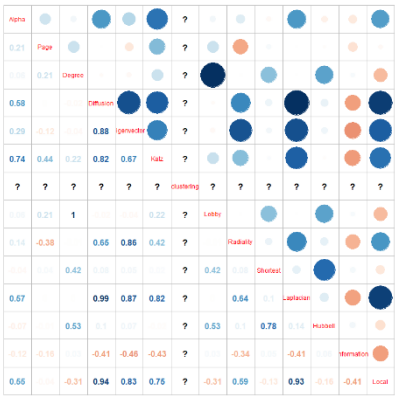
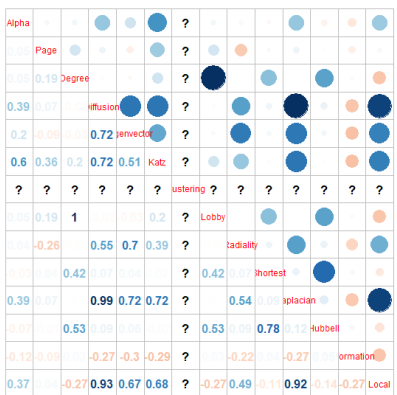
TABLE II. DESCRIPTION OF THE CENTRALITY MEASURES PROPERTIES OF NETWORK DATASETS

S.no.	Centrality Measure	Formulae	Application Areas	Approach used
1.	Alpha Centrality [17]	$\lambda e_i = \sum_j R_{ij} e_j$ R: relationship matrix e: eigenvector	Positive and negative bargaining systems	A total number of paths from a node to other nodes, exponentially attenuated by the length between them. Power comes from being connected to powerless
2.	Page Rank [18]	$PR = (1-d) + \left(d\left(\frac{PR(T_1)}{C(T_1)} + \cdots + \frac{PR(T_n)}{C(T_n)}\right)\right)$ T1...Tn pages C(A): no. of links leaving page A d: damping factor	Ranking webpages by simulating user behavior where most linked and visited webpage is ranked highest	Using the citation graph of hyperlinks in webpages the importance of webpages is derived. A page possesses high page rank if a no. of pages point to it or is pointed by some high page ranked pages
3.	Degree Centrality [19,20]	$\sigma_D(x) = \sum_{i=1}^{n} a_{ix}$ a_{ij}: adjacency matrix $\sigma_D(x)$: centrality score	One of the most basic approach to find a central node. Applicable in almost all the areas where the problem can be generalized in the form of a network and seek to determine important nodes based on their degree	Importance of nodes in a graph depends on its number of links with other nodes i.e. direct friendship connections
4.	Diffusion Centrality [21]	$DC(g;q,T) = [\sum_{t=1}^{T}(qg)^t]$ g: adjacency matrix q: probability T: no. of iterations	Measure the effectiveness of alternative injection points for microfinancing loans.	Approximation of communication centrality. Highly correlated to it and possesses its predictive properties but requires much less data.
5.	Eigen-vector Centrality [22]	$\sigma_E = \frac{1}{\lambda_1}\sum_j A_{ji} v_j$ λ_1: maximum eigenvalue for A A: adjacency matrix v: maximum eigenvector for A	Power measures for a variety of exchange measures	Importance of a node depends on how important are the nodes connected to it
6.	Katz Centrality [23]	$C_K^W(i) = \alpha \sum_j W(i,j).C_K^W(j) + \beta$ W: adjacency matrix α: attenuation constant β: positive constant	Used in a wide variety of directed networks like web network, citation network and biological networks	A generalized form of degree centrality. It is determined by the no. of nodes accessible via a certain path

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

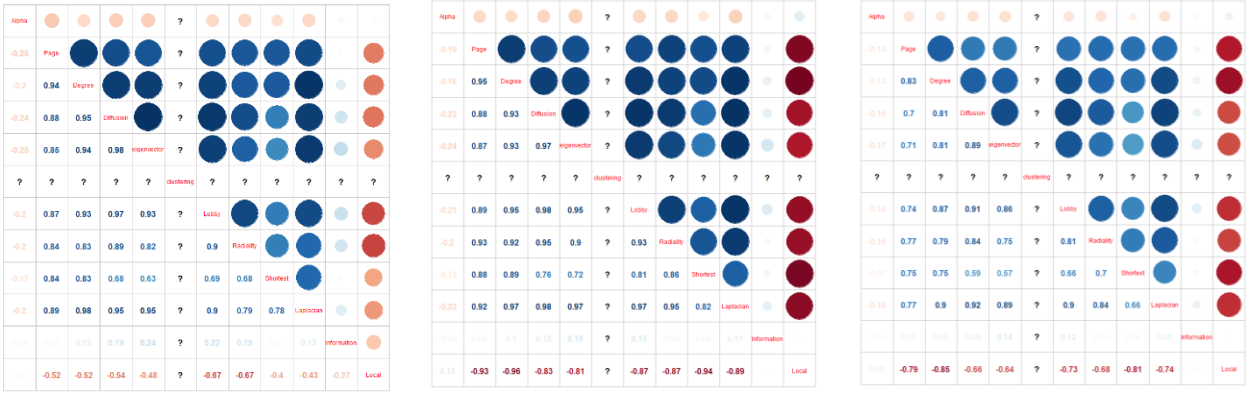
S.no.	Centrality Measure	Formulae	Application Areas	Approach used
7.	Clustering Coefficient [24]	$C(G)= 6n/l$ n: no. of triangles l: no. of paths of length two	Check if a graph is a small world network. Measures how cliquish is the neighborhood of a node	Ratio of existing edges between two nodes within their neighborhood to the total number of possible edges
8.	Lobby index	$l(x) = \max\{k: \deg(y_k) \geq k$ deg(y_k): a degree of neighbor y with at least k degree	Used mainly in scale-free artificial networks	An h-index inspired measure used for general networks. It is equal to the largest no. k for a node n such that n has at least k neighbors with a degree not more than k. Many high degree neighbors mean more influence
9.	Radiality Centrality [26]	$C_R^W(i) = \frac{\sum_{j\in V}(\Delta_G^W + 1 - d^W(i,j))}{g_{jk}^w}$ Δ_G^W:diameter of the graph G $d^W(i,j)$: weighted geodesic distance between nodes i and j	Influential in studying network reachability and connectedness.	Determined by the extent to which a node can access a network as provided by its neighbor nodes. High radiality means lesser time to reach other nodes in the network thus faster sending of information.
10.	Shortest-length Betweenness Centrality [19,20]	$\sigma_B(x) = \sum_{i=1,i\neq x}^{n}\sum_{j=1,j<i,j\neq x}^{n}\frac{g_{ij}(x)}{g_{ij}}$ g_{ij}: shortest path between i and j $g_{ij}(x)$: paths passing through x node	One basic approach to find a central node. Applicable in almost all the areas where the problem can be generalized in the form of a network and seek to determine important nodes based on path length	Communication takes place only along the geodesic paths. Importance of nodes is based on the its relevance for the control and mediation for the information flow
11.	Laplacian Centrality [27]	$E_L(v_i, G) = \frac{E_L(G) - E_L(G_i)}{E_L(G)}$ $E_L(G)$: Laplacian energy with node i $E_L(G_i)$: Laplacian energy without node i	Used for faster results in weighted networks as it has linear computation time.	Defined as the amount of drop seen in the Laplacian energy when a node is removed from the network. Laplacian energy is the sum of the squared eigen values of the Laplacian matrix, obtained by scalar subtraction of weight matrix from the degree matrix of the graph.
12.	H-index [28]	$H_i = max_{1\leq h\leq d_i}\min\left(\left\lvert\mathcal{N}_{\geq h}\right\rvert, h\right)$ $\mathcal{N}_{\geq h}$: neighbor of node i with atleast h degree	The scientific achievement of a scholar is determined by the no. of papers and citation times of those papers measured using h-index. Used for scientometric purposes.	Importance of a node is based on how many other high degree nodes it is connected to. It is equal to the largest no. h for a node n such that n has at least h neighbors with degree not more than h.
13.	Information/ Delta Centrality[29]	$C_i^{\Delta} = \frac{(\Delta P)_i}{P}$ P: cohesiveness measure $(\Delta P)_i$: variation of P after deactivation of i	Useful in finding community structures, characterizing planar graphs and importance of soluble mediators in human immune system	Importance of a node is determined by the response of the network after the node is deactivated. Based on network efficiency and network flow.
14.	Local Bridging Centrality [30]	$C_B(v) = BC(v) \times C_B(v)$ $BC(v) = \frac{d(v)^{-1}}{\sum_{i\in N(v)}\frac{1}{d(i)}}$ BC(v): bridging coefficient d(v): degree of v node N(v): set of node v's neighbors	Identification of key components in scale free networks like biological networks, social networks and web networks	A bridging node is a node connecting two densely connected parts of graph. It's a product of a global measure i.e. betweenness centrality that tells among all the shortest paths how many of them pass through the node and a local measure i.e. bridging coefficient telling how well the node is located between the nodes with high degrees

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

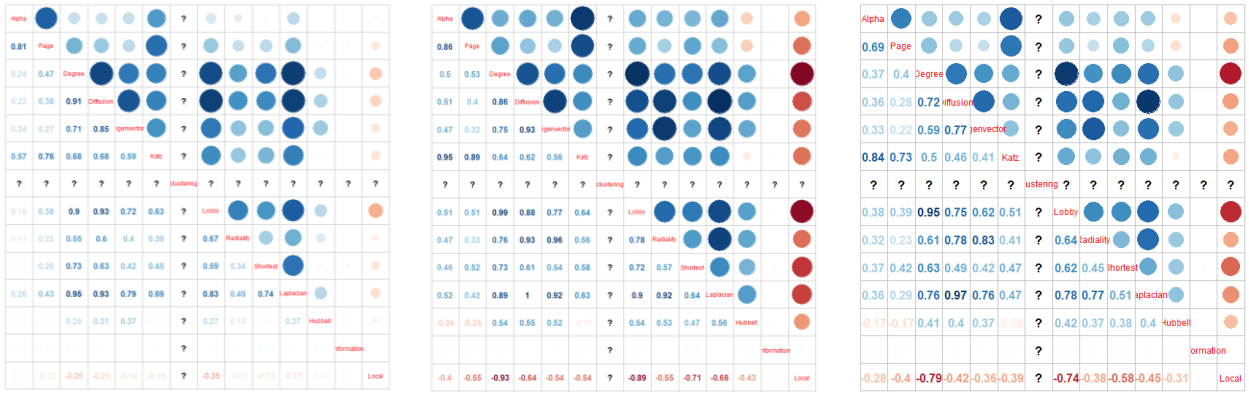
	PEARSON CORRELATION	SPEARMAN CORRELATION	KENDALL CORRELATION
CELEGANS			
DISEASOME			
WIKI-VOTE			
FACEBOOK-EGO			

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

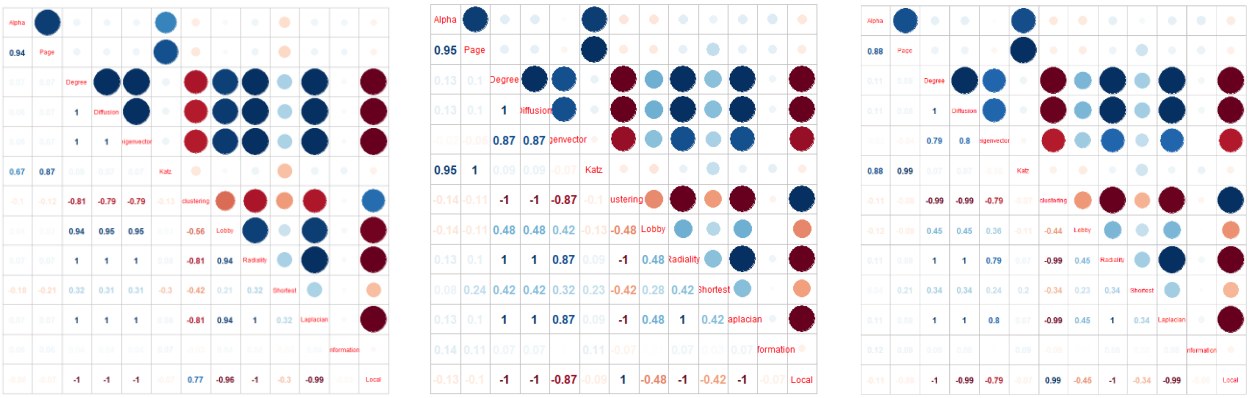
FIRM-HI-TECH



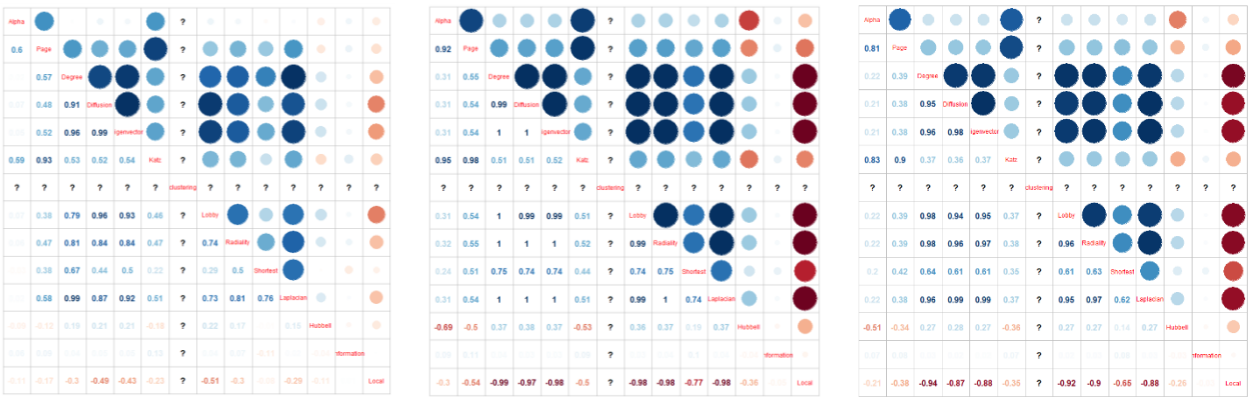
ROUTERS



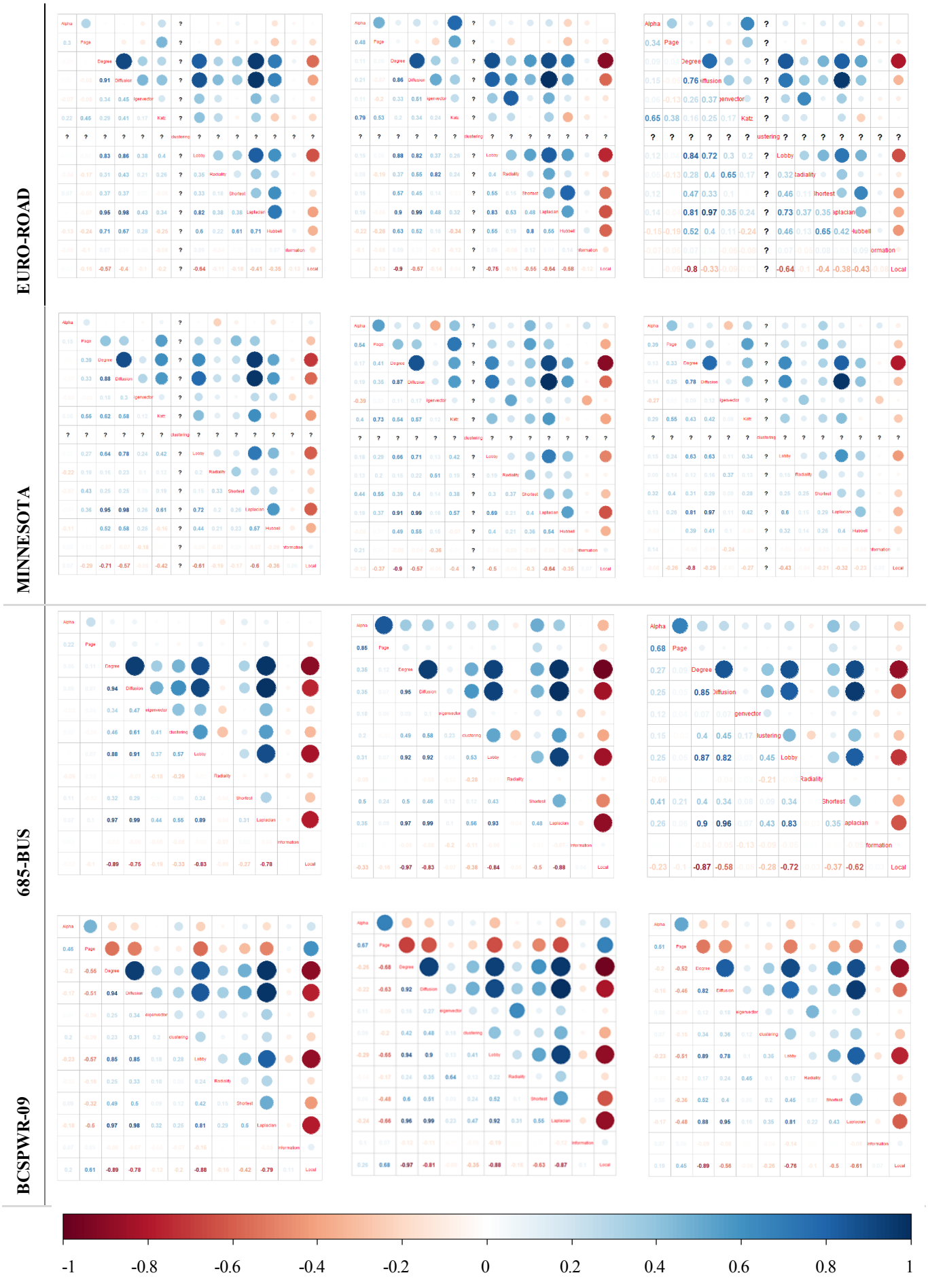
INFECT-HYPER



OCCUPYWALLSTNYC



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

III. RESULTS

Each result is represented in the form of correlation graph as shown in Fig 1. in which the lower triangle shows the numeric values for the respective correlation coefficient obtained for the corresponding pair of centrality measures derived from the row and column. The upper triangle on the other hand shows the visual representation of the results. Blue circles correspond to higher values while red corresponds to the lower values differing from 1 to -1 as shown at the end of the graphs. The size of the circle grows with respect to the size of the higher values. The resultant graphs for all 6 network types are separated by dotted lines.

IV. CONCLUSION

It is important to compare the centrality measures given the fact that there are plenty of such measures while a good result can be provided by a smaller set of them. The comparison of 14 centrality measures across 12 networks displays the variations in results. However, most of the centrality measures have shown strong positive correlations. Only information and local bridging centrality measures possess negative correlation across networks which means the results generated by them differs from other measures are thus good measures to be considered. Clustering centrality has mostly been unable to deliver results except for temporal networks which clearly shows that it is a domain-specific centrality measure. The extent of overlapping information provided by the centrality measures largely depends on network structure. This would reduce the computational time by reducing redundancy. Most centrality measures have moderate to high correlations for all networks with a mean correlation of 90%. Degree and diffusion centrality measures show the highest degree of correlation which further strengthens our proposal that two theoretically varying measures can derive similar results when applied to real-world networks.

REFERENCES

[1] Rossi R. A. and Ahmad N. K. (2015). The network data repository with interactive graph analytics and visualization. *Twenty-Ninth AAAI Conference on Artificial Intelligence* (*AAAI'15*). AAAI Press, 4292–4293.
[2] Batool, K., & Niazi, M. A. (2014). Towards a Methodology for Validation of Centrality Measures in Complex Networks. *PLoS ONE, 9*(4). doi:10.1371/journal.pone.0090283
[3] Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215-239. doi:10.1016/0378-8733(78)90021-7.
[4] Bolland, J. M. (1988). Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social Networks, 10*(3), 233-253. doi:10.1016/0378-8733(88)90014-7
[5] Faust, K. (1997). Centrality in affiliation networks. *Social Networks, 19*(2), 157-191. doi:10.1016/s0378-8733(96)00300-0
[6] Rothenberg, R. B., Potterat, J. J., Woodhouse, D. E., Darrow, W. W., Muth, S. Q., & Klovdahl, A. S. (1995). Choosing a centrality measure: Epidemiologic correlates in the Colorado Springs study of social networks. *Social Networks, 17*(3-4), 273-297. doi:10.1016/0378-8733(95)00267-r
[7] Valente, T. W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How correlated are network centrality measures?. *Connections (Toronto, Ont.)*, *28*(1), 16.
[8] Estrada, E. (2007). Characterization of topological keystone species. *Ecological Complexity, 4*(1-2), 48-57. doi:10.1016/j.ecocom.2007.02.018
[9] Ashtiani, M., Salehzadeh-Yazdi, A., Razaghi-Moghadam, Z., Hennig, H., Wolkenhauer, O., Mirzaie, M., & Jafari, M. (2017). A systematic survey of centrality measures for protein-protein interaction networks. doi:10.1101/149492
[10] Oldham, S., Fulcher, B., Parkes, L., Arnatkevičiūtė, A., Suo, C., & Fornito, A. (2019). Consistency and differences between centrality measures across distinct classes of networks. *Plos One, 14*(7). doi:10.1371/journal.pone.0220061
[11] https://users.dimi.uniud.it/~massimo.franceschet/networks/nexus/examples.html
[12] Latora, V., Crucitti, P., Marchiori, M., & Rapisarda, A. (2003). Complex Systems: Analysis And Models Of Real-World Networks. *Energy and Information Transfer in Biological Systems*. doi:10.1142/9789812705181_0009
[13] Ryan A. Rossi, N. (n.d.). Network Data Repository: The First Interactive Network Data Repository: http://networkrepository.com/
[14] Jalili, M., Salehzadeh-Yazdi, A., Asgari, Y., Arab, S., Yaghmaie, M., Ghavamzadeh, A., & Alimoghaddam, K. (2015, November 16). CentiServer: A Comprehensive Resource, Web-Based Application and R Package for Centrality Analysis. Retrieved August 24, 2020, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4646361/
[15] Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients. *Anesthesia & Analgesia, 126*(5), 1763-1768. doi:10.1213/ane.0000000000002864
[16] Chok, N. S. (2010). *Pearson's versus Spearman's and Kendall's correlation coefficients for continuous data* (Doctoral dissertation, University of Pittsburgh).
[17] Ghosh, R., & Lerman, K. (2012). Rethinking centrality: the role of dynamical processes in social network analysis. *arXiv preprint arXiv:1209.4616*.
[18] Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine.
[19] Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social networks*, *1*(3), 215-239.
[20] Nieminen, J. (1974). On the centrality in a graph. *Scandinavian journal of psychology*, *15*(1), 332-336.
[21] Kang, C., Kraus, S., Molinaro, C., Spezzano, F., & Subrahmanian, V. S. (2016). Diffusion centrality: a paradigm to maximize spread in social networks. *Artificial Intelligence*, *239*, 70-96.
[22] Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social networks*, *29*(4), 555-564.
[23] Zhan, J., Gurung, S., & Parsa, S. P. K. (2017). Identification of top-k nodes in large networks using katz centrality. *Journal of Big Data*, *4*(1), 1-19.
[24] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, *393*(6684), 440-442.
[25] Korn, A., Schubert, A., & Telcs, A. (2009). Lobby index in networks. *Physica A: Statistical Mechanics and its Applications*, *388*(11), 2221-2226.
[26] Valente, T. W., & Foreman, R. K. (1998). Integration and radiality: Measuring the extent of an individual's connectedness and reachability in a network. *Social networks*, *20*(1), 89-105.
[27] Qi, X., Fuller, E., Wu, Q., Wu, Y., & Zhang, C. Q. (2012). Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, *194*, 240-253.
[28] Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, *102*(46), 16569-16572.
[29] Latora, V., & Marchiori, M. (2007). A measure of centrality based on network efficiency. *New Journal of Physics*, *9*(6), 188.
[30] Macker, J. P. (2016, November). An improved local bridging centrality model for distributed network analytics. In *MILCOM 2016-2016 IEEE Military Communications Conference* (pp. 600-605). IEEE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:54:50 UTC from IEEE Xplore. Restrictions apply.

A Deep Language-independent Network to analyze the impact of COVID-19 on the World via Sentiment Analysis

Ashima Yadav, Dinesh Kumar Vishwakarma, *Senior Member, IEEE*

Abstract—Towards the end of 2019, Wuhan experienced an outbreak of novel coronavirus, which soon spread all over the world, resulting in a deadly pandemic that infected millions of people around the globe. The government and public health agencies followed many strategies to counter the fatal virus. However, the virus severely affected the social and economic lives of the people. In this paper, we extract and study the opinion of people from the top five worst affected countries by the virus, namely USA, Brazil, India, Russia, and South Africa. We propose a deep language-independent Multilevel Attention-based Conv-BiGRU network (MACBiG-Net), which includes embedding layer, word-level encoded attention, and sentence-level encoded attention mechanism to extract the positive, negative, and neutral sentiments. The embedding layer encodes the sentence sequence into a real-valued vector. The word-level and sentence-level encoding is performed by a 1D Conv-BiGRU based mechanism, followed by word-level and sentence-level attention, respectively. We further develop a COVID-19 Sentiment Dataset by crawling the tweets from Twitter. Extensive experiments on our proposed dataset demonstrate the effectiveness of the proposed MACBiG-Net. Also, attention-weights visualization and in-depth results analysis shows that the proposed network has effectively captured the sentiments of the people.

Index Terms—Attention, Machine learning, Natural Language Processing, Neural nets, Sentiment analysis.

1 Introduction

THE outbreak of COVID-19 pandemic[1] (popularly referred to as coronavirus), caused by severe acute respiratory syndrome coronavirus 2 (SARS- CoV- 2), has badly hit the entire world. According to the World Health Organization (WHO)[2], this pandemic has affected 7,941,791 people with a death toll of 43,47,96 worldwide. Thus, the WHO declared a Public Health Emergency of International Concern on January 30, 2020. Nearly every country of the world has reported active cases, following which the WHO has advised them to follow strict COVID-19 mitigation guidelines like social distancing, contact tracing, non-essential business closure, quarantine, ban on travel and mass gatherings, etc. [1] [2]. This creates a strong need to assess the attitude and behavior of the people related to the COVID-19 pandemic.

Sentiment analysis is a popular branch in Natural Language Processing which extracts the opinions or views of the people embedded in user-generated data like online reviews, comments, and multimodal data [3] [4]. The field has shown tremendous success in areas like Business review analysis [5] [6], Financial market prediction [7], Recommendation system [8], Multimodal analysis [9]. In the medical domain, the sentiments reflect a patient's health status, which can be good (positive), bad (negative), or normal (neutral) [10]. They can be used for analyzing the drug and medication's effectiveness or confirming the presence or absence of disease [11]. Recently, deep learning models have demonstrated significant improvements in many text processing tasks [12] [13]. They avoid the traditional handcrafted feature engineering by automatically learning the semantic features from the training data itself.

Hence, in this paper, we propose a Multilevel Attention-based Conv-BiGRU Network (MACBiG-Net) for classifying the opinions of the people from the countries which were worse-affected by the COVID-19 pandemic. Firstly, the words are encoded into a real-valued vector with the help of the Glove embedding matrix. Then word-level encoding is performed by 1D CNN, where different convolution kernels help in learning the local features of the text, followed by BiGRU to capture the long term dependencies of a word in both left and right direction. After this, we apply the attention network to extract the most informative word. Finally, we follow the same combination of CNN- BiGRU for encoding the sentence vector and again using attention mechanism to compute the essential weights of the sentences for the entire document classification. In this way, MACBiG-Net can efficiently classify the text into positive, negative, and neutral sentiments. The significant contribution of our work is outlined as follows:

- We propose a deep Multilevel Attention-based Conv-BiGRU Network (MACBiG-Net) for classifying the opinions of the people from the top five worst affected countries by the coronavirus, namely USA, Brazil, India, Russia, and South Africa.
- We also develop the COVID-19 Sentiment dataset by

Ashima Yadav and Dinesh Kumar Vishwakarma are with Biometric Research Laboratory, Department of Information Technology, Delhi Technological University, Delhi, India. E-mail: ashimayadavdtu@gmail.com, dvishwakarma@gmail.com.

[1]https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

[2]https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200815-covid-19-sitrep-208.pdf?sfvrsn=9dc4e959_2 (As on 15 August, 2020)

crawling and downloading the tweets from the famous microblogging site, Twitter, and labelling them according to the sentiments belonging to the positive, negative, and neutral sentiment classes. We download and analyze the tweets of the people belonging to the USA, Brazil, India, Russia, and South Africa, starting from January 1, 2020, to June 7, 2020.

- The proposed MACBiG-Net is language-independent as it can classify the tweets belonging to multiple languages like Hindi, Japanese, Arabic, Spanish, Urdu, etc. effectively by encoding the words and sentences, and assigning larger weights to relevant contexts for the document classification.
- We conduct rigorous experiments on the proposed COVID-19 Sentiment dataset. Experimental results demonstrate that the proposed network show significant improvement over the popular baseline methods for sentiment classification in terms of Precision, Recall, F1 score, Accuracy, Receiver operating characteristic (ROC), and Area under the curve (AUC) evaluation metrics. We also visualize the attention weights of the network for getting in-depth knowledge about the attention mechanism.
- Finally, we draw several crucial observations by visualizing the country-wise sentiments of the people from January to June. We also analyze the popular positive and negative sentiment topics, words, and hashtags that were generated during the pandemic to see how the pandemic affected the lives of people all over the world. This analysis can serve as the feedback to the government agencies regarding the mitigation plans taken by them. Further, it may also guide the future planning of the public health agencies in case of any such outbreak.

The remainder of the paper is organized as follows: Section 2 discusses the Related Work for sentiment analysis in the health-care domain. Section 3 focuses on the proposed methodology by describing the proposed MACBiG-Net. Section 4 presents the experimental results by discussing quantitative and qualitative analysis. Section 5 concludes our paper and also focuses on future work.

2 Related Work

The sentiment analysis is a vast field having ample applications, as discussed above. Since we could not find any previous work which had focused on sentiments expressed during the COVID-19 pandemic, hence we discuss prominent works related to sentiments in the health-care domain only. The facets of sentiments in medicine are broadly categorized into [10]: capturing patient's health status (good/bad/no change), diagnosing the certainty or severity of a disease, and the effectiveness of the treatment or drug.

The early work in this area focuses on in-built lexicons or machine learning-based approaches for extracting the sentiments. Yang et al. [14] proposed a framework based on the weighted scheme Latent Dirichlet Allocation (LDA) to analyze the health-related information regarding breast cancer. The medical topics were clustered together to perform sentiment analysis using AFINN lexicon. Each topic was associated with positive and negative sentiment terms. The framework can benefit other users to seek advice or take necessary precautions in case they experience similar symptoms. Sabra et al. [15] proposed a framework based on semantic extraction and sentiment assessment of the risk factors to predict the diagnosis of Venous thromboembolism disorder. The clinical narratives were parsed to extract the risk factors with the help of MetaMap. These risk factors were stored in a dictionary, and semantic-rich rules were derived from weighting the risk factors according to the severity. Finally, the polarity of the adjectives and adverbs were extracted with the help of SentiWordNet lexicon. The framework achieved 54.5% precision and 85.7% recall to identify patient's likelihood to develop VTE early and start the treatment as soon as possible.

Several clinical trials are necessary before releasing any drug to the market. Still, discovering all the side effects of the drugs is a challenging task. Moh et al. [16] developed an architecture to study the side effects of five crucial drugs by capturing the tweets from Twitter. The sentiment scores were extracted using several lexicons like SentiWordNet, AFINN, etc. Similarly, Jiménez-Zafra et al. [17] applied machine learning classifiers and lexicon-based approaches to compute the sentiment polarities for the physicians and the drugs. The reviews related to physicians contain informal language, while discussions about drugs contain the combination of colloquial language and specific terms like adverse effects or name of the drug. Korkontzelos et al. [11] identified the adverse drug reactions on social media platforms by integrating the sentiment feature. Experimental results show that sentiment features improved the F-score from 72.14% to 73.22%.

Recently, deep based approaches are being explored to extract the health-care sentiments expressed by the people. Limsopatham et al. [18] applied CNN and word embeddings on social media messages to medical concepts. Chen et al. [19] extracted the emotion features based on the LSTM network for screening the users on perinatal depression. The results show that the method improves the early screening process for perinatal depression. Grisstte et al. . [20] applied the Bi-LSTM model for identifying the adverse drug reactions from social media text. Talpada et al. [21] studied the impact of Telemedicine for heart attacks and epilepsy. They compared the lexical and semantic methods with the deep based methods. The lexical and semantic methods were based on VADER and TextBlob, whereas LSTM was employed for deep based classification. Experimental results signify that sentiments can be used to obtain the demographic distribution of Telemedicine for hearth attacks and epilepsy. Thus, sentiment in the context of medicine shows the impact of a disease or drug on the life of a person over a period of time. This can provide feedback to assess the effectiveness of the treatment or measures taken to combat a disease or a crucial event.

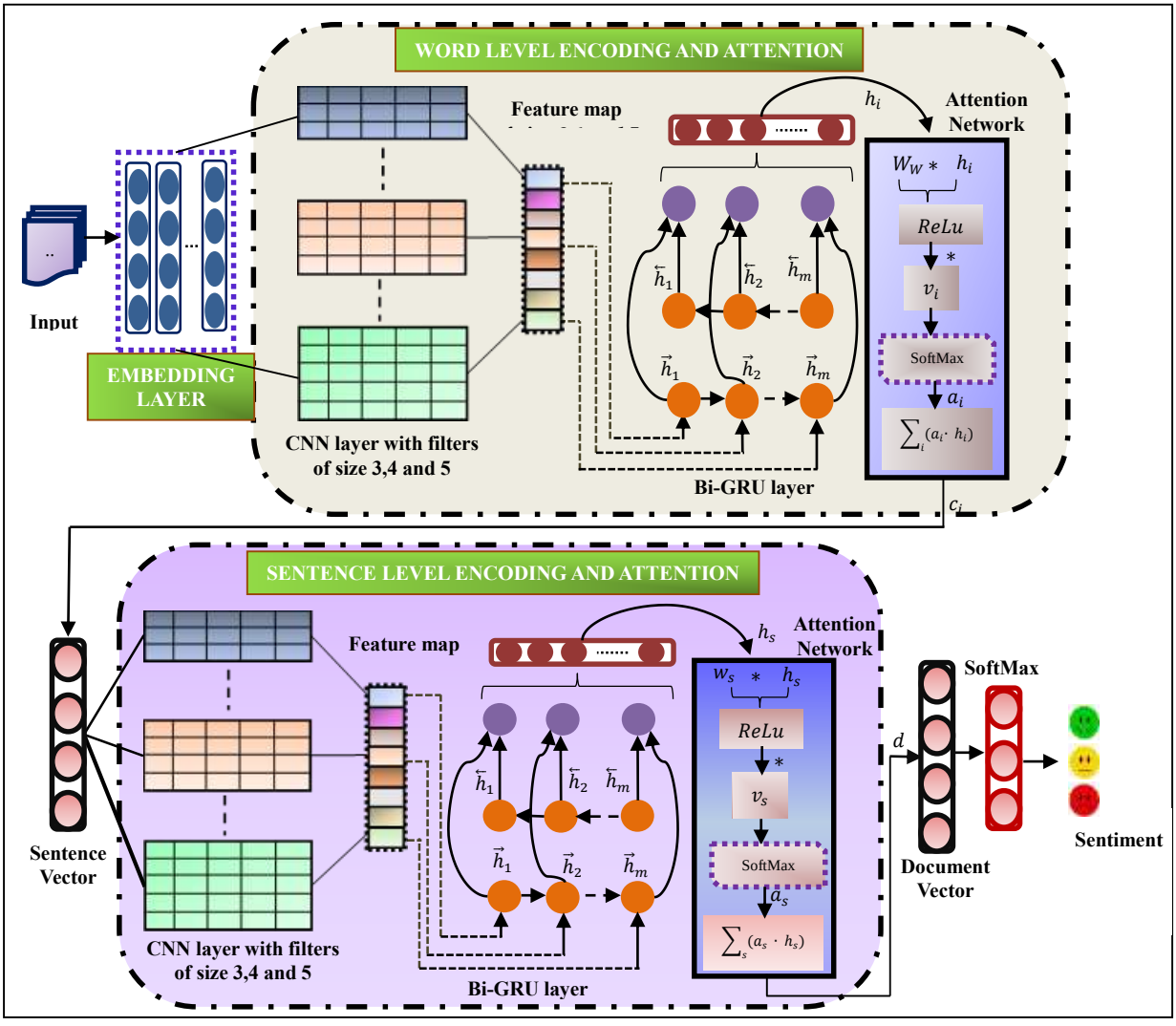


Fig. 1. Block diagram of the proposed Multilevel Attention-based Conv-BiGRU Network (MACBiG-Net)

3 Proposed Methodology

In this section, we explain the proposed framework for sentiment classification. It describes the data preprocessing process where we follow a two-step data cleaning approach, followed by the proposed Multilevel Attention-based Conv-BiGRU network (MACBiG-Net) to learn the semantic information of a sentence by encoding the essential words and sentences in a sequence

3.1 Data preprocessing

We scrapped the data from one of the most popular microblogging site Twitter for all the five countries based on their location. The complete data scraping, cleaning, and labeling process are discussed in Section 4. The final dataset comprises of people's views belonging to three output classes: positive, negative, and neutral. Since data processing is one of the most crucial tasks for handling the unstructured textual data, hence we meticulously designed a two-step data cleaning approach. The first approach is followed before translating the data into the English language so that the translation process is not affected by special characters or non-ASCII codes in the sentence. This includes removing hyperlink, special characters/symbols, retweet headers, twitter usernames, non-ASCII codes, NaN (Not a Number) character, and whitespaces. Finally, all the uppercase words are converted into lowercase by using the NumPy library functions.

In the second approach, the labeled dataset goes through stemming and lemmatization to reduce the word to its root form. Porter Stemmer and WordNet Lemmatizer are used for stemming and lemmatization, respectively. We also removed the stop words that frequently occur in the text and convey no meaningful information, followed by hashtags and punctuations removal. The cleaned data is then passed to the Multilevel Attention-based Conv-GRU network, as discussed in the subsequent section for the final classification.

3.2 Multilevel Attention-based Conv-BiGRU Network (MACBiG-Net)

The section describes the proposed MACBiG-Net, as shown in Fig. 1. The network includes three main steps: Embedding Layer, which encodes the input words into the low-dimensional vectors, Word-level Encoding and

Table I
Layer description of Step 2: Word-level Encoding and Attention

Layer Name	Input Shape	Output Shape	# Parameters	Connected to
Input Layer	(None, 200)	(None, 200)	0	-
Embedding	(None, 200)	(None, 200, 100)	1835200	Input Layer
Conv1D_1	(None, 200, 100)	(None, 198, 128)	38528	Embedding
Conv1D_2	(None, 200, 100)	(None, 197, 128)	51328	Embedding
Conv1D_3	(None, 200, 100)	(None, 196, 128)	64128	Embedding
MaxPooling1D_1	(None, 198, 128)	(None, 66, 128)	0	Conv1D_1
MaxPooling1D_2	(None, 197, 128)	(None, 65, 128)	0	Conv1D_2
MaxPooling1D_3	(None, 196, 128)	(None, 65, 128)	0	Conv1D_3
Concatenate	(None, 66, 128), (None, 65, 128), (None, 65, 128)	(None, 196, 128)	0	MaxPooling1D_1, MaxPooling1D_2, MaxPooling1D_3
MaxPooling1D_4	(None, 196, 128)	(None, 65, 128)	0	Concatenate
Bidirectional_GRU	(None, 65, 128)	(None, 65, 200)	183200	MaxPooling1D_4
TimeDistributed (Dense)	(None, 65, 200)	(None, 65, 100)	20100	Bidirectional_GRU
Hierarchical_attention_Network	(None, 65, 100)	(None, 100)	10200	TimeDistributed (Dense)

Attention, in which the words of a sentence are encoded by 1D Conv-BiGRU based representation to obtain the sentence representation, followed by word-level attention which computes the essential weights for the sentence vector, and Sentence level encoding and attention, where the sentences of a document are again encoded by 1D Conv-BiGRU based representation to obtain the document representation, followed by sentence-level attention to get the document vector for the classification.

3.2.1 Embedding Layer

Each word in the sequence of the sentence is converted into a real-valued vector. Formally, let $\{r_1, r_2, \ldots\ldots, r_n\}$ be the sequence of n words in a sentence. We first embed each word w_i into a real-valued vector through a word embedding matrix $E \in R^{d_r * |V|}$, where d_r is the word embedding size, and V is the vocabulary size. We use the Glove[3] embedding matrix to get the final sequence of vector $\{e_1, e_2, \ldots\ldots, e_n\}$, which serves as an input for the next layer.

3.2.2 Word-level Encoding and attention

Our proposed architecture is inspired by [22], which intends to capture the structure of words from the sentence and the structure of sentences from a document. The word encoding is performed by 1D convolution, followed by the Bi-GRU mechanism. We first describe the convolution process as follows:

The local features are extracted by using a 1D CNN network with different kernel sizes to generate a feature map. The different convolution kernels help in learning the various local characteristics of the text. In general, for e_i^{th} embedding corresponding to r_i^{th} word, the concatenation operation $\oplus$ is expressed as follows:

$$e_{1:n} = e_1 \oplus e_2 \oplus \ldots\ldots\ldots \oplus e_n \quad (1)$$

The convolution operation is applied on the window of j words to generate a new feature value m_i as shown below:

$$m_i = f\,(W \cdot e_{i:i+j-1} + b) \quad (2)$$

Where, b = bias and f = Rectified Linear unit activation function (ReLu). This filter is applied to different possible window of words in a sentence which generates the following feature map:

$$M = [m_1, m_2, \ldots\ldots, m_{n-j+1}] \quad (3)$$

We apply the filter of sizes 3, 4, and 5 to obtain multiple features followed by max-pooling layer to capture the maximum feature value. The concatenated output of all three filter sizes is passed into the BiGRU based layer, which contains the hidden states to store crucial information. The bidirectional nature of GRU is based on the fact that words in a sentence are not only related to the previous words but also to the following words. We have chosen GRU over LSTM due to its efficiency in training, which makes it computationally cheaper than LSTM. The forward GRU with $\vec{h}_i = \overrightarrow{GRU}\left(e_i, \vec{h}_{i-1}\right)$is concatenated with backward GRU, $\overleftarrow{h}_i = \overleftarrow{GRU}\ (e_i, \overleftarrow{h}_{i+1})$to obtain the output of BiGRU at step w as $h_i = [\vec{h}_i, \overleftarrow{h}_i]$. Thus, the BiGRU layer fetches the contextual information of every sentence centered around the word w_i. This makes sense as the orientation of sentiment needs to consider the past and future context information in a sentence. Thus, the word is encoded by the Conv-BiGRU block by storing the local and contextual features.

The attention mechanism is based on the idea that when users read a document word by word in each sentence, then they would pay attention to the most informative word or sentence. Hence, we model both the word level and sentence level attention differently. In word-level attention, the word annotation h_i is fed into the network to get its vector representation v_i. We apply the *ReLu* activation function over *tanh* as specified in [22], as *ReLu* converges quickly and results in cheaper computation. This is described in the (4):

$$v_i = ReLu\ (W_w \cdot\ h_i + b_w) \quad (4)$$

[3] https://nlp.stanford.edu/projects/glove/

Table II
Layer description of Step 3: Sentence-level Encoding and Attention

Layer Name	Input Shape	Output Shape	# Parameters	Connected to
Input Layer	(None, 15, 200)	(None, 15, 200)	0	-
TimeDistributed (Model)	(None, 15, 200)	(None, 15, 100)	2202684	Input Layer
Conv1D_1	(None, 15, 100)	(None, 13, 128)	38528	TimeDistributed (Model)
Conv1D_2	(None, 15, 100)	(None, 12, 128)	51328	TimeDistributed (Model)
Conv1D_3	(None, 15, 100)	(None, 11, 128)	64128	TimeDistributed (Model)
MaxPooling1D_1	(None, 13, 128)	(None, 4, 128)	0	Conv1D_1
MaxPooling1D_2	(None, 12, 128)	(None, 4, 128)	0	Conv1D_2
MaxPooling1D_3	(None, 11, 128)	(None, 3, 128)	0	Conv1D_3
Concatenate	(None, 4, 128), (None, 4, 128), (None, 3, 128)	(None, 11, 128)	0	MaxPooling1D_1, MaxPooling1D_2, MaxPooling1D_3
MaxPooling1D_4	(None, 11, 128)	(None, 3, 128)	0	Concatenate
Bidirectional_GRU	(None, 3, 128)	(None, 3, 200)	183200	MaxPooling1D_4
TimeDistributed (Dense)	(None, 3, 200)	(None, 3, 100)	20100	Bidirectional_GRU
Hierarchical_attention_Network	(None, 3, 100)	(None, 100)	10200	TimeDistributed (Dense)
Dropout	(None, 100)	(None, 100)	0	Hierarchical_attention_Network
Dense	(None, 100)	(None, 3)	303	Dropout

The resulting annotation v_i is multiplied (dot product) by the word context vector v_{wt} which is learned during the training. The softmax function is applied to obtain attention weights a_i as shown in (5) below:

$$a_i = \frac{\exp(v_{wt} \cdot v_i)}{\sum_i \exp(v_{wt} \cdot v_i)} \quad (5)$$

Finally, the attention weights a_i is concatenated with the word annotations to obtain the sentence vector c_i as described in (6) below:

$$c_i = \sum_i (a_i \cdot h_i) \quad (6)$$

The layer details of the word-level encoding and attention step are shown in Table I.

3.2.3 Sentence level encoding and attention

As discussed, selecting important sentences in the document becomes another crucial task for the classification. Hence, we follow the same procedure for generating the document representation of the sentences and, as discussed in Section 3.2.2, by encoding the sentences with relevant context and computing the crucial weights of these contexts for the document classification.

The sentence vector c_i obtained in Section 3.2.2 is passed into the 1D CNN layer and Bi-GRU layer to encode the sentence and obtain information h_s of the neighboring sentences centered around sentence s. The sentence context vector v_{st} is used to compute the sentence level attention a_s, which further yields the document vector d that summarizes all the information of the sentences in a document. Mathematically, this is represented as below in (7)-(8).

$$v_s = ReLu\ (W_s \cdot h_s + b_s) \quad (7)$$

$$a_s = \frac{\exp(v_{st} \cdot v_s)}{\sum_s \exp(v_{st} \cdot v_s)} \quad (8)$$

$$d = \sum_s (a_s \cdot h_s) \quad (9)$$

The layer details of the Sentence level encoding and attention step are shown in Table II. The learned vector d of [1* 100] dimension is passed to the dropout layer with a 0.5 rate to handle the overfitting problem in deep neural network architectures. This is followed by a dense layer and softmax activation function. The entire network is trained end to end for obtaining the best weights by monitoring the validation accuracy of the model. We use categorical cross-entropy loss, which is defined in (10) as follows:

$$Loss = -\sum_{j=1}^{3} (y_j \cdot \log \hat{y}_j) \quad (10)$$

Where, y_j is the target value corresponding to the model output $\hat{y}_j$ for the j^{th} sample. We have used the L2-regularization parameter, which penalizes the larger weights in the network for avoiding the high variance problem. This makes the objective function of the network as follows:

$$Cost\ function = Loss + (\lambda/2m * \sum ||w||^2) \quad (11)$$

Here, λ is the regularization parameter and one of the hyperparameters of the network, which is tuned according to the validation set. Our model achieves optimized results with λ = 0.001.

4 Experimental Analysis

This section discusses the data collection and labeling process in Section 4.1, implementation details in Section 4.2. We evaluate the effectiveness of the network in Section 4.3 and perform baseline comparison in Section 4.4. Finally, we visualize our classification results and extract the sentiment topics in Section 4.5.

4.1 Dataset Collection and Labeling

In order to analyze the impact of the novel coronavirus, we extracted the Tweets from January 1, 2020 to June 7, 2020 of top five most affected countries, namely, USA (52,03,206 cases), Brazil (32,24,876 cases), India (25,26,192 cases), Russia (9,17,884 cases), and South Africa (5,79,140 cases) as per

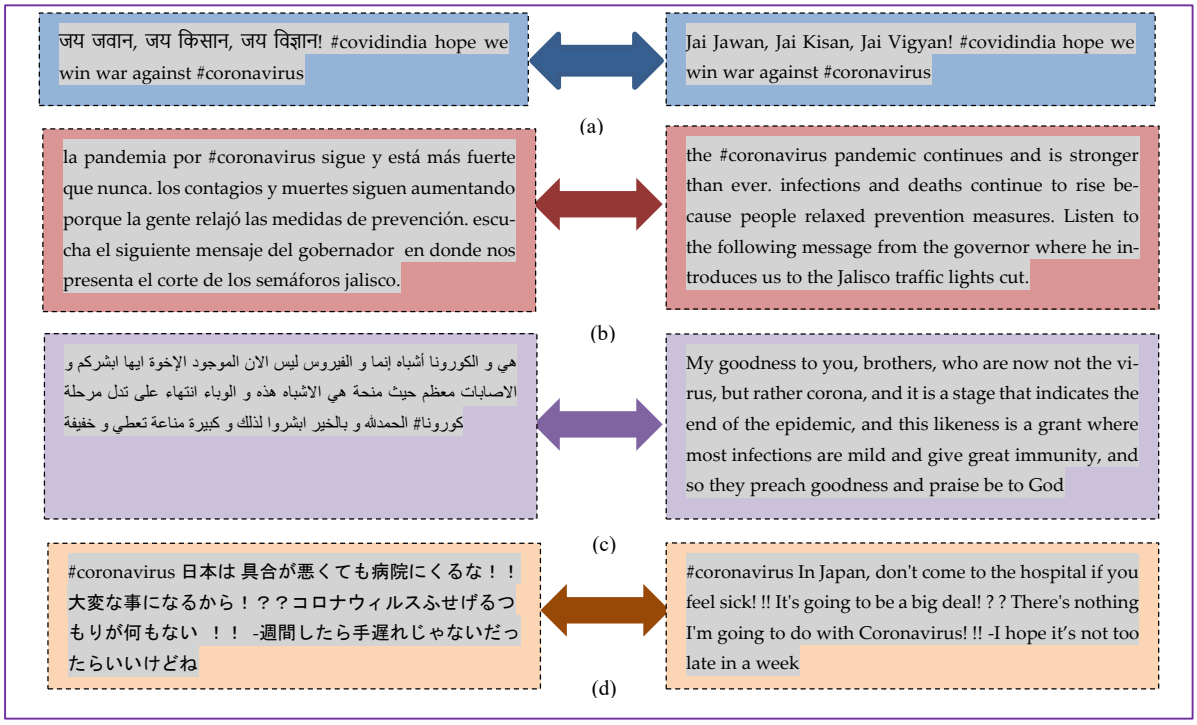


Fig. 2. Sample tweets in different languages from COVID-19 Sentiment Dataset (a) Hindi to English (b) Spanish to English (c) Arabic to English (d) Japanese to English.

data available from WHO[4]. We used some popular hashtags like #coronavirus, #covid19, and #COVID_19. The scrapped tweets were preprocessed as discussed in Section 3.1 and translated into English by the Google translator. Some sample tweets in a different language and their English translation is shown in Fig. 2.

Table III
Country-wise details of COVID -19 Sentiment Dataset

Country	Sentiment			Total
	Positive	Negative	Neutral	
USA	383	321	431	1135
Brazil	267	318	444	1029
India	302	164	299	765
Russia	147	168	249	564
South Africa	200	193	232	625
Total	**1299**	**1164**	**1655**	**4118**

We used the Textblob[5] library to weakly label the processed tweets with the sentiment score of [-1, 1]. The tweets with +1 sentiment scores were labeled as positive, tweets with -1 score were labeled as negative, and finally, tweets with 0 sentiment scores were labeled as neutral. However, we observe that in many cases, the Textblob library fails to classify the data into correct sentiment labels. E.g., Consider the following tweet, "*My dad tested positive for the corona*" was labeled as positive sentiment, though it belongs to the negative sentiment category. Hence, the shortlisted tweets were again manually annotated (strong labels) by three humans with positive (+1), negative (-1), and neutral (0) sentiment category. Finally, 4118 tweets were annotated with an average value of Cohen Kappa inter-annotator agreement as 0.85. The country-wise dataset details corresponding to each sentiment category are shown in Table III.

4.2 Implementation Details

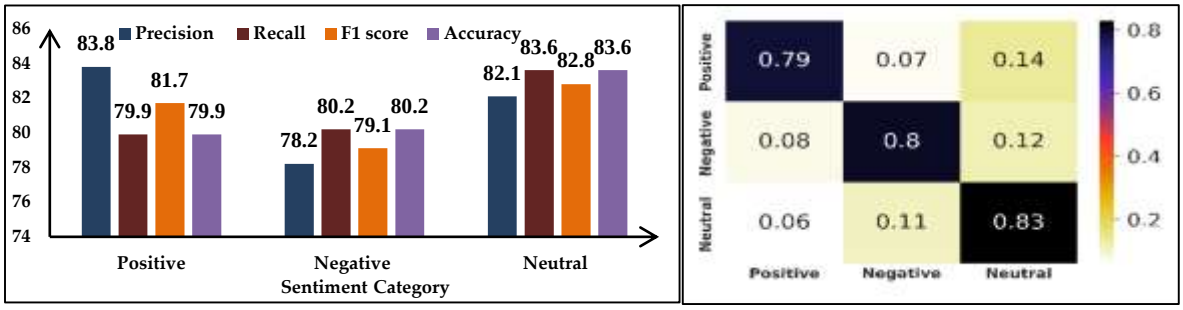
The proposed MACBiG-Net is implemented on Python 3 using the popular Keras framework. The experiments were performed on Windows 10 machine with 128GB RAM using NVIDIA Titan RTX GPUs. The embedding dimension is 100, maximum length of all sequences is 200, and Punkt sentence Tokenizer is used for dividing the text into a list of sentences. Adam optimizer with learning rate = 0.0001, default parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$ is set for training. The entire network is trained end to end with 64 batches for 100 epochs. Dropout and Regularization are used for avoiding overfitting. We have used a ten-fold cross-validation testing strategy with 80% data for training, 5% data for validation, and 15% data for testing. The validation accuracy is monitored during training the network. The model achieving the highest validation accuracy is selected for the testing phase.

4.3 Classification Results and Observations

We compare the sentiment prediction performance of the MACBiG-Net in terms of precision, recall, F1 score, and accuracy corresponding to different sentiment categories, as shown in Fig. 3 (a). Since we used a ten-fold cross-validation strategy, the final results are reported by averaging the results across each of the test fold.

The accuracy of positive, negative, and neutral sentiment category is 79.9%, 80.2%, and 83.6%, respectively.

[4] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200815-covid-19-sitrep-208.pdf?sfvrsn=9dc4e959_2 (As on 15 August, 2020)
[5] https://textblob.readthedocs.io/en/dev/



(a) (b)

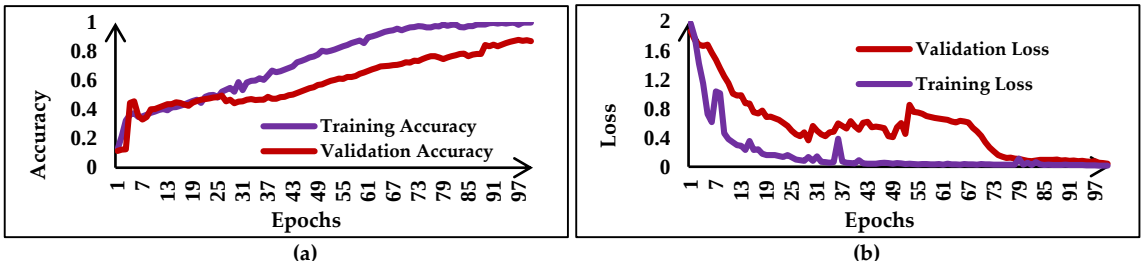
Fig. 3. (a) Sentiment classification results on COVID-19 Sentiment Dataset (%) (b) Confusion Matrix.

The average accuracy of the network is 81.5%. Further, the confusion matrix in Fig. 3 (b) gives an in-depth analysis of the classification results. It can be seen that a major amount of misclassification occurs when a positive sample is incorrectly classified as neutral. Yet, we can see that the network can discriminate well across each of the sentiment categories. These results clearly explain the effectiveness of MACBiG-Net.

We also provide accuracy and loss curves for training and validation set in Fig. 4 (a) and (b), respectively. As we can see, the training accuracy is 100%, which means the model is trained completely, and the validation accuracy shows how effectively the model can perform on the unseen samples. Similarly, decreasing loss curves with each epoch confirm the adequate learning of the model and validate its internal classification performance.

The foremost aim of our work is to analyze the sentiments of people expressed globally on the COVID-19 pandemic. This can be visualized in Fig. 5, where we plot the positive (+1), negative (-1), and neutral (0) sentiments of people over time located in the top five worst affected countries by the pandemic. As discussed, we retrieved the tweets from January 1, 2020 to June 7, 2020. However, in many countries, sentiments were generated in the month of late February or March. This may be because the first confirmed case in that country started in late February (Eg: *The virus was first confirmed in Brazil on February 25, 2020* [6]), or no specific sentiments were generated during January and February. A five-day window represents the time component in Fig. 5 along the x-axis. This means that each bar represents the sentiment expressed in these five days. However, each day, several sentiments were being generated. Hence, we took the sentiment that was voiced the maximum number of times by the people on any particular day as the final sentiment for that day. We summarize our observations as follows:

- In the USA, significant sentiments emerged from March onwards. This may be because Trump declared a national emergency[7] on March 13, and by the end of March, the cases were reported in all 50 US States[8]. In early March, positive sentiment was recorded. However, by the end of March till mid-April, negative sentiment was prevalent. Similarly, by the end of May, we observed mostly negative sentiments. One of the primary reasons was the breaking news of George Floyd, which created a state of panic and terror in the USA from May 25 onwards. People started tweeting that COVID-19 will continue to spread as citizens were protesting and moving in groups, without following any preventive measures. Overall, it was observed that 43.6% of sentiments were neutral, 37.9% were positive, and 18.3% were negative.
- In India, positive sentiments were observed majorly in March and April. This may be because people were hopeful in the initial stages to overcome the virus. Further, Prime Minister (PM) Narender Singh Modi came up with some ideas to boost the support of all the frontline workers. On March 22, the PM requested every Indian to clap their hands or ring some bells from their house to boost the morale of frontline workers. Similarly, on April 5, the PM urged everyone to light candles or *diyas* for nine minutes as a mark to fight against the deadly virus. Several people followed these tasks and happily enjoyed them. Towards the end of May, we see the people expressed negative sentiments. Since, by the end of May, a large number of active cases were



(a) (b)

Fig. 4. Tracing internal performance of MACBiG-Net with (a) Accuracy curve for train and validation set (b) Loss curves for train and validation set.

[6] https://www.gov.br/saude/pt-br
[7]https://edition.cnn.com/2020/03/13/politics/donald-trump-emergency/index.html
[8]https://archive.vn/20200428073138/https://banningca.gov/DocumentCenter/View/7139/CDC_COVID19-Weekly-Key_Messages_03292020_FINAL

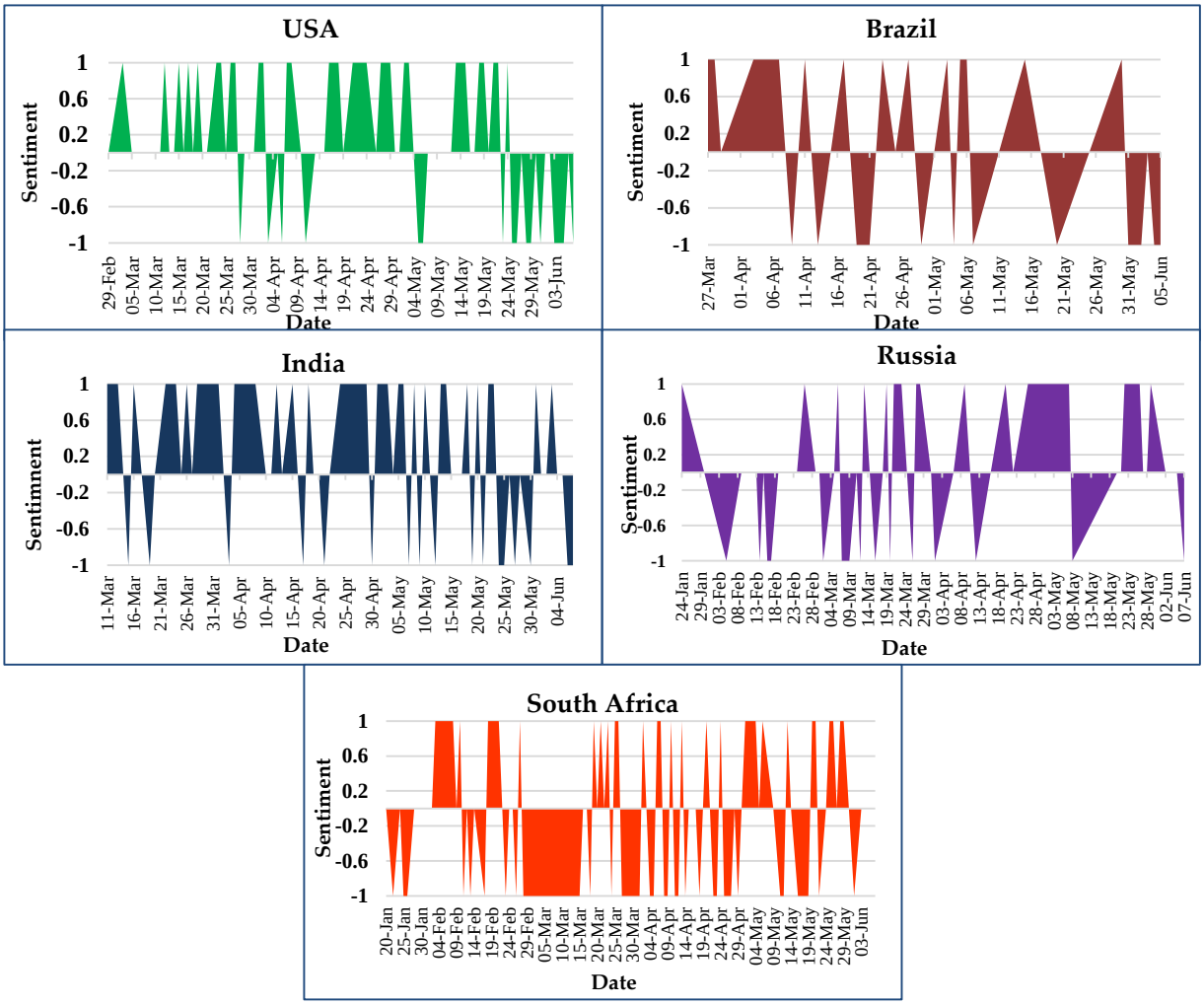


Fig. 5. Visualizing the sentiments of people over time in top five worst affected countries by COVID-19 pandemic.

being reported each day. Overall, 50.6% of positive sentiments, 28.3% of neutral sentiments, and 20.9% of negative sentiments were observed in India.

- Russia reported its first positive case on January 31, 2020. People have majorly shown neutral and negative sentiments during the mid of February. However, positive sentiments were expressed in April. Overall, 39.2% of neutral sentiment, 33.9% of positive sentiment, and 26.7% of negative sentiment were observed in Russia.
- In South Africa, majorly negative sentiments were being observed from late February till mid of March. People finally realized that the deadly virus had hit their country as the number of cases started rising sharply. However, in April, we can see some positive sentiments as citizens started motivating everyone by asking them to wear masks, read books, involve in daily exercise, and posting positive ideas by their former President Nelson Mandela. Overall, South Africa expressed 39.6% of negative sentiments, 32.6% of positive sentiments, and 27.7% of neutral sentiments.
- Finally, we conclude that overall, 38.3% of positive sentiments, 32.6% of neutral sentiment, and 28.3% of negative sentiments were expressed by these top five affected countries.

4.4 Baseline Comparison

We use several popular baseline methods for comparison with the proposed MACBiG-Net. The details are discussed as follows:

Table IV
Comparative Results of different methods on COVID-19 Sentiment Dataset (%)

Method	Accuracy	Precision	Recall	F1 score
LSTM [23]	72.6	73.3	73.6	73.4
Bi-LSTM [24]	73.2	74.9	72.0	73.4
CNN [25]	75.8	73.1	75.2	74.1
CNN-RNN [26]	77.7	78.6	74.2	76.3
HAN [22]	79.4	77.4	78.9	78.1
Self-attention-based LSTM [27]	70.9	68.2	66.7	67.4
RNN-Capsule [28]	71.5	68.5	70.8	69.6
BERT [29]	76.6	75.2	74.8	74.9
MACBiG-Net (Ours)	**81.5**	**81.3**	**81.2**	**81.2**

- **LSTM** [23]**:** The classical long-short memory network is applied. Word vectors are initialized by Glove. The hidden layer size is 300 units. Adam optimizer with learning rate of 0.001 is used.
- **Bi-LSTM** [24]**:** This represents standard bi-directional LSTM. We follow the same settings as described above for LSTM.
- **CNN** [25]**:** Convolutional neural network with multiple filters of size [3,4,5], each having 100 filters is applied. Adam optimizer is used with a learning rate of 0.001.
- **CNN-RNN** [26]**:** We successively stack the CNN layer with max-pooling layer followed by RNN units. For RNN, we experimented with LSTM and GRU variants. The GRU based integration shows the highest performance.
- **HAN** [22]**:** The hierarchical attention network uses two LSTM layer followed by attention mechanism at word-level and sentence-level for composing the final text representation. The HAN is the simplest form of our proposed approach as it does not encounter localized features.
- **Self-attention based LSTM** [27]**:** The self-attention mechanism is used with Bi-LSTM, which learns the contribution of each hidden state by providing a set of summarized weight vectors for each hidden state. A learning rate of 0.06 is used.
- **RNN-Capsule** [28]**:** We apply capsule networks where the input instance representation is taken from the LSTM unit. The dimension of hidden vectors is set to 256. The attention mechanism is used to construct the capsule representation inside another capsule. All word vectors are initialized with Glove. Adam optimizer is used with a learning rate of 0.001.
- **BERT** [29]: We use a pre-trained Bidirectional Encoder Representation from Transformers (BERT) model with the help of an online released TensorFlow[9] library. The transformer blocks are set to 12, hidden size to 768, and self-attention head to 12.

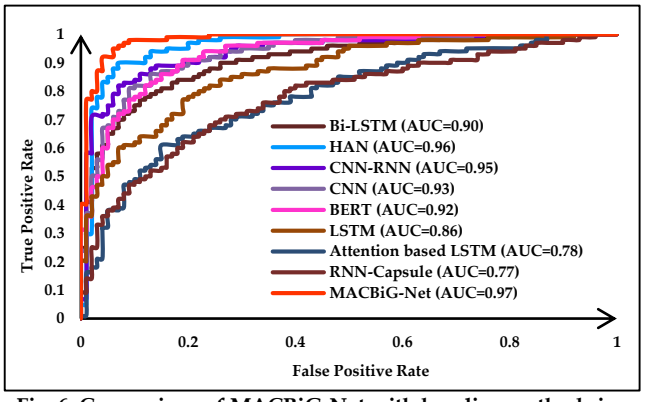


Fig. 6. Comparison of MACBiG-Net with baseline methods in terms of ROC curve and Area under the curve (AUC).

As evident from Table IV, the proposed network has outperformed the baseline methods and recent works in terms of sentiment classification by achieving 81.5% average accuracy. However, the HAN [22] and CNN-RNN [26] provides good enough results for the classification. This serves as the primary motivation of our proposed MACBiG-Net, which inherits the basic idea of both the methods and achieves nearly 2% and 4% higher accuracy compared to [22] and [26], respectively. This can also be visualized from the ROC curves and area under the curve (AUC), shown in Fig. 6, where we plot the ROC curves for the baseline methods. The AUC values help in comparing the ROC curves in a better way. From the AUC values, it is clear that MACBiG-Net has consistently outperformed all the previous methods for sentiment classification.

4.5 Qualitative Analysis

In this section, qualitative analysis is performed to visualize the word-level attention and extract the positive and negative sentiment topics that were generated during the pandemic.

4.5.1 Visualizing Word-level Attention

In this section, qualitative analysis is performed on MACBiG-Net, to investigate the role of word-level attention for the sentiment classification. We randomly picked four tweets from our dataset, two of them were positive, and two were negative. Thus, in Fig. 7, (a) – (b) belongs to the positive sentiment class (shown in Red color), and (c) – (d) belongs to the negative class (shown in Blue color). The color signifies the attention weights of each word. The darker the color in the respective category, the higher amount of attention is given to the word.

We can see that attention focuses on the prominent part of the sentence, by giving darker color to the most important word. E.g., In review sentences (a) - (b), positive sentiment is conveyed by the crucial words like *'brilliant'*, *'happy'*, *'healthy'*, *'safety'*, *'good'*, which have received more score by the attention mechanism for better sentiment prediction. Similarly, in reviews (c)-(d), negative sentiment is reflected by words like *'scary'*, *'pandemic'*, *'death'*, which again have received more attention score. The words which convey no informative meaning (like stop words), were given very low weights by highlighting them with light colors. This shows that the attention weights focus on those words which have more relation with the output sentiment category. Thus, the visualization shows how the attention weights are changing along with the words in a sentence.

4.5.2 Visualizing positive and negative sentiment topics

To get an in-depth analysis on the opinion of the people, we apply LDA (Latent Dirichlet Allocation) based topic modeling to determine the set of words (topics) that contribute to generate the positive and negative sentiments. This will help in finding the hidden thematic structure of the document and drawing essential conclusions from them. We use LDAvis [30] tool for visualization, which further helps in finding the prevalent topics in a document, relation between the topics, and the conditional distribution of a given term over the set of topics. The optimal number of topics was determined by building many LDA models with a different number of topics (k) and selecting

[9] https://github.com/tensorflow/tensor2tensor

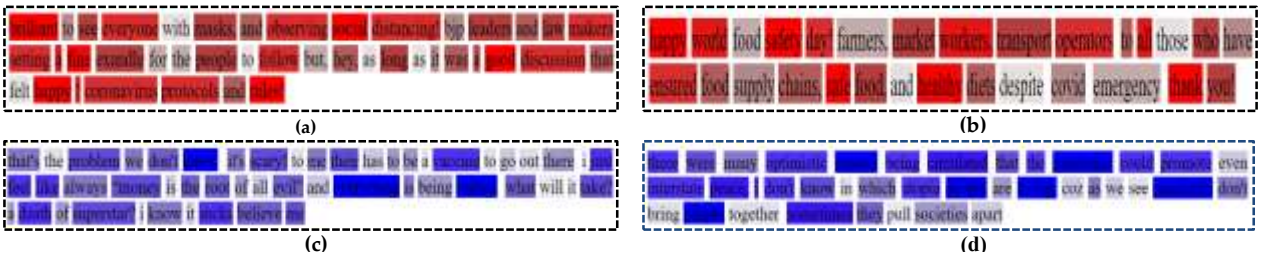


Fig. 7. Visualization the word-level attention weights from proposed dataset: (a)-(b) Positive Sentiment (c)-(d) Negative Sentiment.

the one with the highest coherence score [31]. This gives us k =10. The layout of LDAvis is shown in Fig. 8 (a)-(b) for positive and negative sentiment labels, respectively.

If we consider (a), the entire layout is divided into two panels: the left panel shows the ten topics in the form of circles where the similar topics are plotted close to each other, the right panel shows the bargraphs which represent the terms or words that are most useful for understanding the currently selected topic (from the left panel). The length of the gray bars depicts the entire document-wise frequencies of each term, and the red bars depict the topic-wise frequencies of each term. Thus, for a given topic, we can compare the length of the red and gray bars to understand the relevance of the topic. Further, selecting a term on the right panel tells the conditional distribution of the selected term over the topics (on the left panel). In this way, topic modeling helps in explaining the underlying information in a document. Hence, we summarize our observations as follows:

- The prevalent positive words related to COVID-19 on twitter are: 'good', 'stay safe', 'happy', 'best', 'thank you', 'support', 'doctors', 'mask', 'together', 'proud' and the popular negative words linked to COVID-19 on twitter are: 'virus', 'deaths', 'bad', 'trump', 'help', 'lost', 'fake', 'died', 'outbreak', 'infected', 'hate', 'worst', 'crisis', 'fight'.
- Many positive tweets were directed towards the frontline workers where users were motivating and boosting the morale of the health care workers by appreciating them, donating them masks, dedicating some songs, providing free accommodation to them, and donating money to restaurants for distributing meals to essential workers.
- In some countries like Brazil, health care workers celebrate the discharge of the first COVID-19 patient from the hospital. In India, many citizens were delighted to see everyone unified with candles or flashlights in their hands from their balconies.
- Finally, positive sentiments were directed towards maintaining a healthy diet, reuniting with families after travel restrictions were lifted, boosting the immune system, asking everyone to stay safe, binge-watching television series, encouraging everyone to take precautions, wishing and motivating the users on special occasions by posting positive messages on occasions like Birthdays, Mothers Day, Earth Day, World Bicycle Day, Memorial Day (celebrated in the US), World Labor Day, International Nurses Day, and staying together to beat the deadly virus.
- During the pandemic, several social-economic factors also contributed to generate negative sentiments in people all around the world. In the US, the negative tweets were mainly focused on the racial injustice related to the death of George Floyd, which resulted in massive protests and violence even during the time of the pandemic. People fear that these mass demonstrations may occur in an uptick in the COVID-19 cases.
- In India, Mumbai (one of the worst-hit cities by COVID-19), was severely affected by cyclone winds. In West Bengal, Cyclone Amphan caused extensive damage. The poor, jobless migrants went traveling back to their home by walking thousands of kilometers due to lockdown restrictions in India.
- In general, many people also lost their jobs. The unemployment rate was increased during the pandemic. People were bored and anxious during home-quarantine. The upsurge in infected cases also created fear and stress among people. Further, lots of fake news[10] was generated in social media, creating panic about the disease.
- The popular positive hashtags are: #salutetocoronafighters, #inthistogether, #lovemycountry, #withme, #thankyouindia, #indiafightscorona, #flattenthecurve, #fightthevirus, and #frontlineheroes and negative hashtags are: #DictatorTrump, #BlackLivesMatter, #GeorgeFloyd, #AmericaInCrisis, #Riot2020, #migrantlivesmatter, #pandemia, #chinavirus, #wuhanvirus, and #fuck_coronavirus.

5 Conclusion

This paper is aimed to analyze the impact of the novel coronavirus on the people by extracting and analyzing the sentiments of the top five most affected countries with the virus. We propose MACBiG-Net, which aims to extract the positive, negative, and neutral sentiment using word-level encoded attention and sentence level encoded attention. This is achieved by developing and labeling the COVID-19 Sentiment Dataset, which contains 4118 labeled tweets crawled from Twitter. Further, we visualize the sentiments of the people from January 31 to June 7 and summarize essential observations. We conclude that during the pandemic, lots of sentiments were being generated, and people were posting their views or opinions on several topics like frontline workers, upsurge in active cases, travel restrictions, and about the virus itself. Much fake news was also generated during this time, which was misleading the

[10] https://timesofindia.indiatimes.com/blogs/toi-editorials/social-media-menace-beware-of-fake-news-going-viral-faster-than-covid-19-it-will-cost-lives-literally/

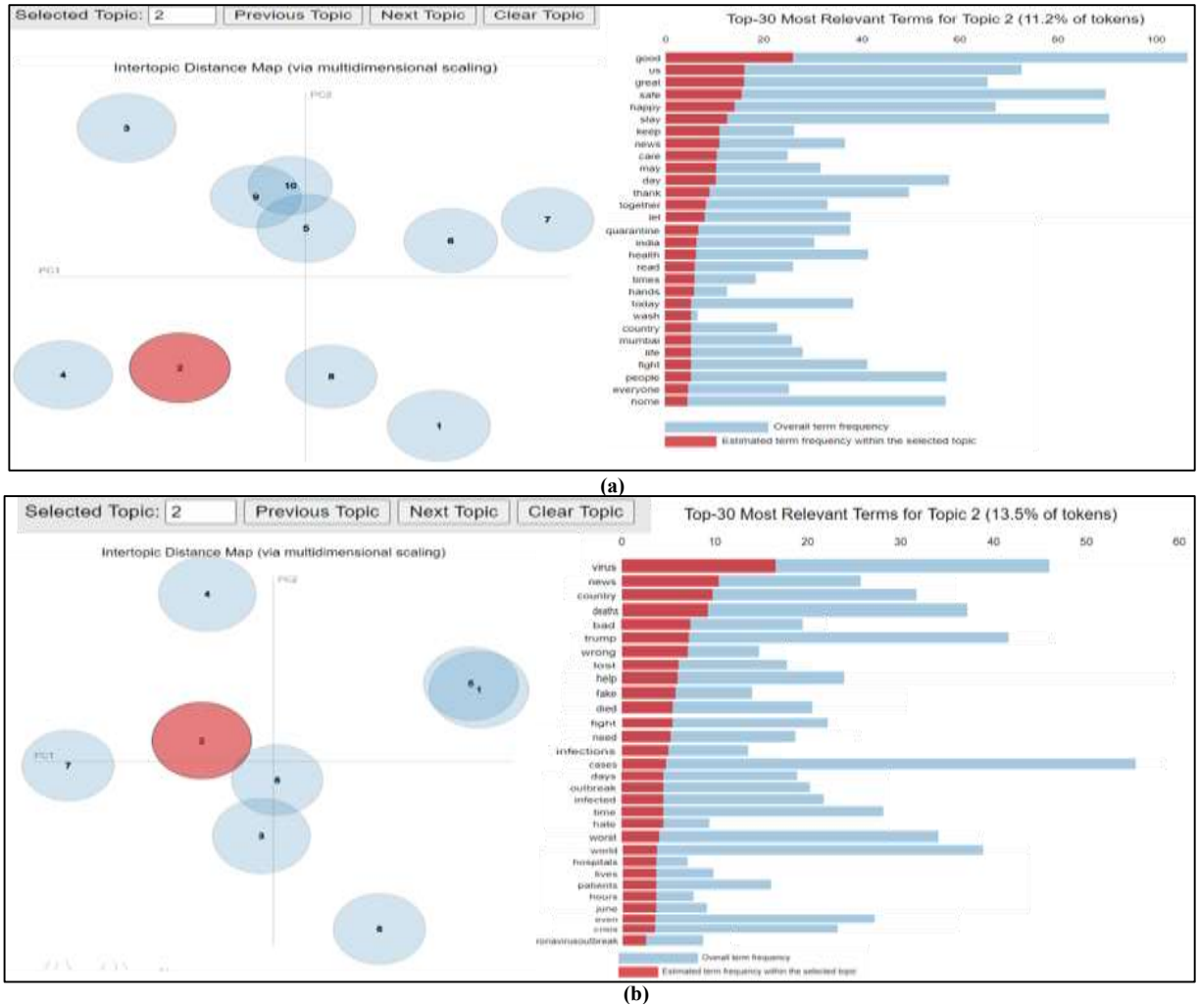


Fig. 8. Identifying the hidden information in document through LDAvis visualization for (a) Positive and (b) Negative sentiment

people. However, despite being stressed, people appreciated the efforts of all frontline workers and motivated each other to follow all precautions.

In the future, we aim to expand our study to analyze the impact of the virus on the rest of the world by doing in-depth research on several strategies like home-quarantine, work from home, social distancing undertaken to fight the deadly virus. We also aim to study the sentiments generated by other popular nations in the world.

REFERENCES

[1] R. M. Anderson, H. Heesterbeek, D. Klinkenberg and T. D. Hollingsworth, "How will country-based mitigation measures influence the course of the COVID-19 epidemic?," *The Lancet,* vol. 395, no. 10228, pp. 931-934, 2020.

[2] J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, T. W. Russell, J. D. Munday, A. J. Kucharski and W. J. Edmunds, "Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts," *The Lancet Global Health,* vol. 8, no. 4, pp. e488-e496, 2020.

[3] A. Yadav and D. K. Vishwakarma, "A comparative study on bio-inspired algorithms for sentiment analysis," *Cluster Computing,* p. 1–21, 2020.

[4] M. Tsytsarau and T. Palpanas, "Managing Diverse Sentiments at Large Scale," *IEEE Transactions on Knowledge and Data Engineering,* vol. 28, no. 11, pp. 3028-3040, 2016.

[5] S. M. Mohammad, P. Sobhani and S. Kiritchenko, "Stance and sentiment in tweets," *ACM Transactions on Internet Technology (TOIT),* vol. 17, no. 3, pp. 1-23, 2017.

[6] W. Zhao, Z. Guan, L. Chen, X. He, D. Cai, B. Wang and Q. Wang, "Weakly-supervised deep embedding for product review sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering,* vol. 30, no. 1, pp. 185-197, 2017.

[7] T.-c. Huang , R. N. Zaeem and K. S. Barber, "It Is an Equal Failing to Trust Everybody and to Trust Nobody: Stock Price Prediction Using Trust Filters and Enhanced User Sentiment on Twitter," *ACM Transactions on Internet Technology (TOIT),* vol. 19, no. 4, pp. 1-20, 2019.

[8] C. Huang, W. Jiang, J. Wu and G. Wang, "Personalized Review Recommendation based on Users' Aspect Sentiment," *ACM Transactions on Internet Technology (TOIT),* vol. 20, no. 4, pp. 1-26, 2020.

[9] A. Yadav and D. K. Vishwakarma, "A deep learning architecture of RA-DLNet for visual sentiment analysis," *Multimedia Systems,* vol. 26, p. 431–451, 2020.

[10] K. Denecke and Y. Deng, "Sentiment analysis in medical settings: New opportunities and challenges," *Artificial Intelligence in Medicine,* vol. 64, pp. 17-27, 2015.

[11] I. Korkontzelos, A. Nikfarjam, M. Shardlow, A. Sarker, S. Ananiadou and G. H. Gonzalez, "Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts," *Journal of Biomedical Informatics,* vol. 62, pp. 148-158, 2016.

[12] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review,* vol. 53, no. 6, pp. 4335-4385, 2019.

[13] Y. Wang, Q. Chen, M. Ahmed, Z. Li, W. Pan and H. Liu, "Joint Inference for Aspect-level Sentiment Analysis by Deep Neural Networks and Linguistic Hints," *IEEE Transactions on Knowledge and Data Engineering,* pp. 1-12, 2019.

[14] F.-C. Yang, A. J. Lee and &. S.-C. Kuo, "Mining Health Social Media with Sentiment Analysis," *Journal of medical systems,* vol. 40, no. 11, pp. 1-8, 2016.

[15] S. Sabra, K. M. Malik and M. Alobaidi, "Prediction of venous thromboembolism using semantic and sentiment analyses of clinical narratives," *Computers in Biology and Medicine,* vol. 94, pp. 1-10, 2018.

[16] M. Moh, T. Moh, . Y. Peng and L. Wu, "On adverse drug event extractions using twitter sentiment analysis," *Network Modeling Analysis in Health Informatics and Bioinformatics,* vol. 6, no. 1, pp. 1-12, 2017.

[17] S. M. Jiménez-Zafra, M. T. Martín-Valdivia, M. D. Molina-González and L. A. Ureña-López, "How do we talk about doctors and drugs? Sentiment analysis in forums expressing opinions for medical domain," *Artificial intelligence in medicine,* vol. 93, pp. 50-57, 2019.

[18] N. Limsopatham and N. Collier, "Normalising Medical Concepts in Social Media Texts by Learning Semantic Representation," *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics,* p. 1014–1023, 2016.

[19] Y. Chen, B. Zhou, W. Zhang, G. Gong and G. Sun, "Sentiment Analysis Based on Deep Learning and Its Application in Screening for Perinatal Depression," *IEEE Third International Conference on Data Science in Cyberspace,* pp. 451-456, 2018.

[20] H. Grisstte and E. Nfaoui, "Daily life patients Sentiment Analysis model based on well-encoded embedding vocabulary for related-medication text," *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining,* pp. 921-928, 2019.

[21] H. Talpada, M. N. Halgamuge and N. T. Q. Vinh, "An Analysis on Use of Deep Learning and Lexical-Semantic Based Sentiment Analysis Method on Twitter Data to Understand the Demographic Trend of Telemedicine," *11th International Conference on Knowledge and Systems Engineering (KSE), Vietnam,* pp. 1-9, 2019.

[22] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola and E. Hovy, "Hierarchical attention networks for document classification," *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, San Diego,* pp. 1480-1489, 2016.

[23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation,* vol. 9, no. 8, pp. 1735-1780, 1997.

[24] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural networks,* vol. 18, no. 5-6, pp. 602-610, 2005.

[25] Y. Kim, "Convolutional neural networks for sentence classification," *Empirical Methods in Natural Language Processing,* pp. 1746-1751, 2014.

[26] S. Lai, L. Xu, K. Liu and J. Zhao, "Recurrent convolutional neural networks for text classification," *Twenty-ninth AAAI conference on artificial intelligence,* p. 2267–2273, 2015.

[27] Z. Lin, M. Feng, C. Nogueira, M. Yu, B. Xiang, B. Zhou and Y. Bengio, "A structured self-attentive sentence embedding," *International Conference on Learning Representations,* 2017.

[28] Y. Wang, A. Sun, J. Han, Y. Liu and X. Zhu, "Sentiment Analysis by Capsules," *Proceedings of the 2018 world wide web conference,* pp. 1165-1174, 2018.

[29] J. Devlin, M.-W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT),* p. 4171–4186, 2019.

[30] C. Sievert and K. S. Shirley, "LDAvis: A method for visualizing and interpreting topics," *Proceedings of the workshop on interactive language learning, visualization, and interfaces,* pp. 63-70, 2014.

[31] K. Stevens, P. Kegelmeyer, D. Andrzejewski and D. Buttler, "Exploring topic coherence over many models and many topics," *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning,* pp. 952-961, 2012.

Ashima Yadav received B.Sc.(Hons) in Computer Science from University of Delhi, New Delhi, India in 2013, and M.C.A. from Guru Gobind Singh Indraprastha University, New Delhi, India in the year 2016. She is currently working towards the Ph.D. degree from the Department of Information Technology, Delhi Technological University, New Delhi, India. Her current research interest includes deep learning, natural language processing, machine learning, Emotion Recognition, and sentiment analysis.

Dinesh Kumar Vishwakarma (M'16, SM'19) received the B.Tech. degree from Dr. Ram Manohar Lohia Avadh University, Faizabad, India, in 2002, the M.Tech. degree from the Motilal Nehru National Institute of Technology, Allahabad, India, in 2005, and the Ph.D. degree degree in the field of Computer Vision and Machine Learning from Delhi Technological University University (Formerly Delhi College of Engineering), New Delhi, India, in 2016. He is currently an Associate Professor with the Department of Information Technology, Delhi Technological University, New Delhi. His current research interests include Computer Vision, Machine Learning, Deep Learning, Sentiment Analysis, Fake News and Rumour Analysis, Crowd Behaviour Analysis, Person Re-Identification, Human Action and Activity Recognition. He is a reviewer of various Journals/Transactions of IEEE, Elsevier, and Springer. He has been awarded with "Research Excellence Award" by Delhi Technological University, Delhi, India in 2018, 2019 and 2020.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
DVD Part Number:CFP20OSV-DVD; ISBN: 978-1-7281-5463-3

A Study on Combating Emerging Threat of Deepfake Weaponization

Rahul Katarya
Department of Computer Science and Engineering,
Delhi Technological University
New Delhi, India
rahulkatarya@dtu.ac.in

Anushka Lal
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
anushkalal_2k18co083@dtu.ac.in

Abstract—A breakthrough in the emerging use of machine learning and deep learning is the concept of autoencoders and GAN (Generative Adversarial Networks), architectures that can generate believable synthetic content called deepfakes. The threat lies when these low-tech doctored images, videos, and audios blur the line between fake and genuine content and are used as weapons to cause damage to an unprecedented degree. This paper presents a survey of the underlying technology of deepfakes and methods proposed for their detection. Based on a detailed study of all the proposed models of detection, this paper presents SSTNet as the best model to date, that uses spatial, temporal, and steganalysis for detection. The threat posed by document and signature forgery, which is yet to be explored by researchers, has also been highlighted in this paper. This paper concludes with the discussion of research directions in this field and the development of more robust techniques to deal with the increasing threats surrounding deepfake technology.

Index Terms—Deep Learning, Generative Adversarial Networks, autoencoders, Deepfake detection, Fake image, Fake Video

I. INTRODUCTION

Deepfake technology is a new automatic computer graphics tool that portrays entirely unrealistic events as real through digital media manipulation. Deepfake gained its name from the Reddit platform where an anonymous user used this term, a combination of deep learning and fake, for replacing celebrities into adult video clips. As soon as the code was made public, widespread interest spawned in the users about the generation of fake content. ObamaNet [1], was an architecture that featured an impressive use of lip-syncing technology to generate synchronized photo-realistic lip-sync videos. Deepfake technique makes it possible to generate unauthentic videos of people expressing or saying things they have never said before [2].

Deepfake makes use of Artificial Intelligence (AI), machine learning, and deep learning concepts. AI deals with intelligence at the machine level.

Machine learning is an evolving concept of Computer Science where machines can be trained to learn from provided data and accordingly take decisions on their own, just like humans do. Deep learning is a broader aspect of machine learning, where highly complex networks are trained to learn from a massive database of unstructured data.

Today there are various free deepfake applications like the Chinese app, Zao [3] that lets users to easily swap faces with movie stars so they can see their self, playing that role in the movie, DeepNude [4], that can create nonconsensual porn, FakeApp, FaceSwap, and DeepFace Lab. The existence of such open-source software and the availability of devices in the market for fabricating and propagating these falsified information has brought to attention the immediate need for detection and elimination of malicious deepfake content. Deepfakes can act as a powerful weapon to insurgent groups and terrorist organizations, who may depict their adversaries using inflammatory words or engaging in provocative actions, to maximize the galvanizing impact on their target audiences. For instance, a member of the Islamic State (or ISIS), can falsely generate fake content that shows government officials or soldiers discussing bombing attacks at a mosque, to aid their terrorist group's recruitment [5]. States can use this weapon to undermine their non-state opponents. Deepfakes can affect the outcome of an election, hence a threat to democracy. Deepfake is even weaponizing satellite images of Earth by showing the existence of certain objects in landscapes and locations that do not exist in reality, just so they can play with the minds of military analysts and influence their decisions based on these fake images [6][7].

Amidst the threats posed by deepfakes in various sectors, the ability to generate realistic simulations can have a whole new positive impact on humanity. It can create an array of opportunities in fields of education, entertainment, and business. Historical figures can be made to communicate with students. In movies, face-swapping can be achieved for scenes

that cannot be fulfilled by the actors alone. For example, in 2016's Rogue One, late Peter Cushing's appearance as Grand Moff Tarkin was possible through similar technology [8]. Deepfakes can be used in business such that customers can exactly view their appearance in products they wish to buy without applying them in reality. In the medical world, this technology can play a great role in training doctors, nurses, and surgeons to operate on real-life scenarios in a virtual environment [9].

However, the potential of deepfakes to cause a broad spectrum of serious harm to society is a matter of greater concern. They can act as a new weapon to humiliation and destruction, identity theft and exploitation, defamation, and manipulation of legal evidence. Several methods have been proposed to detect deepfakes by focusing on the minute details of the content such as facial texture, head poses, eye-blinking, skin color, lip movements, Spatio-temporal features, and capsule forensics. Most of these rely on the same deep learning techniques that are used for the creation of deepfakes.

In the foreseeable future, deepfakes will continue evolving; thus, it is important to investigate their development and improve the methods of detection accordingly. The main objective of this paper is to present a survey of methods used for the creation and detection of deepfakes. Section II explains the popular underlying principle of deepfake architecture. Section III discusses and compares different proposed methods for deepfake detection. Section IV presents the contribution of this paper to the survey. The research opportunities in this direction are further highlighted in Section V.

II. DEEPFAKE CREATION

Deepfakes use Deep Neural Networks (DNNs). DNNs consists of a set of interconnected units called neurons. These units together perform some form of computational task and help solve complex problems. Two popular technologies associated with deepfake creation are the autoencoder-decoder model and the GAN architecture that has been discussed below.

A. Autoencoders

Autoencoder was the first technology to be used in deepfake creation. Autoencoder is used to recreate images that it is trained on. The output generated works in three different phases: encoder, latent space, and a decoder [10]. The encoder first compresses the input pixels to a relatively smaller size by encoding special attributes like skin texture, skin color, facial expressions, open eye, closed eye, head pose, and any minute details of the face. This compressed image is sent as input to latent space that is useful for understanding and learning patterns and structural similarities between the data points [11]. Lastly, the decoder decompresses this information to reconstruct an output based on its representation in latent space. The decoder tries to recreate an image that resembles the original as much as possible.

Fig. 1 shows how an autoencoder can be used to swap two faces. Following the path indicated by the red arrows, Face B is reconstructed similar to Face A. The important aspect here is that both the faces have used the same encoder. This will help the encoder to use general features that are common to both faces, and their positioning in latent space will also be similar. This will allow the autoencoders to transform the same picture with faces of the target and the original individual swapped. Here, the latent space of Face A is referred for Decoder of Face B to be able to reconstruct Face B similar to Face A. This technique has its application in various deepfake technologies like DFaker, DeepFaceLab, and TensorFlow-based deepfakes [12].

B. *Generative Adversarial Networks*

The majority of the current deepfake technologies incorporate the use of GAN. The GAN architecture was first proposed by Ian Goodfellow in 2014 [13]. He introduced a framework that involves the use of two neural networks that work by challenging each other: the first one generates new data while the other discriminates this new data from the original training data set. Contesting the two neural networks tend to improve both the quality of fake data produced as well as the neural network's discrimination ability. If a large number of images are fed to GAN, it can create a unique image on its own [14]. However, it is necessary to attach a filter that can help differentiate these unique outputs as acceptable or not. For this, GANs make use of a discriminative network that checks the generated data with true data. Both are trained to operate together until the discriminator has falsely classified the generated output as authentic, almost 50% of the time. This helps us conclude that the generator model is successfully generating plausible examples. Fig.1 shows the block diagram explaining the workflow of GAN architecture.

The GAN architecture uses the min-max method for training the generator and discriminator [13]. The min (0) represents a fake output while the max (1) represents a genuine output. The goal of the discriminator is to get as close as possible to the max value such that a realistic-looking deepfake is generated which can be further be used for face-swapping in images and videos. GANs are more suitable for generating new data [15]. The main advantage of GANs over autoencoders is that they can be used for a wider range of tasks for example to produce several classes of data, similar to the MNIST dataset [16]. Autoencoders, on the other hand, are more suitable for compressing data to lower dimensions or generating semantic vectors from it.

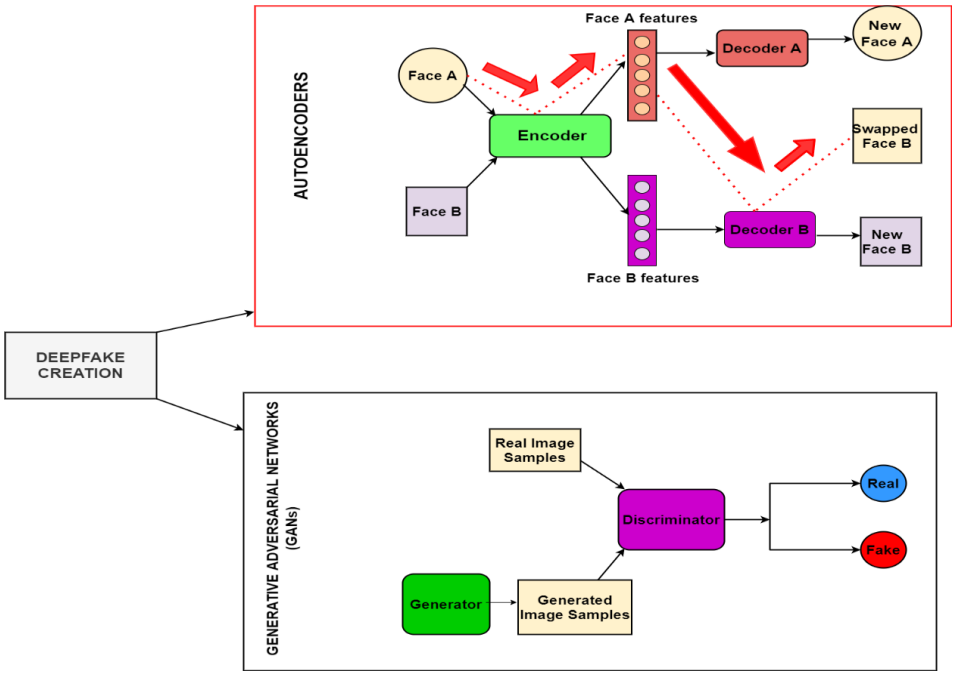


Fig. 1. The workflow of Autoencoders and GAN in the creation of Deepfakes

B. DEEPFAKE DETECTION

To date, there is much ongoing research in this field. According to data obtained from https://app.dimensions.ai, there have been a total of 931 research publications in this field by the end of August 2020 and is likely to increase further in the coming years. Image, video, and audio are the three major types of deepfakes that are currently being dealt with globally.

Malicious use of deepfakes has become undeniably powerful. It has become a personal weapon of revenge. They are used to manipulate financial markets or to destabilize international relations. Several efforts have been made to prevent against this deceptive and destructive potential of deepfakes. This section presents a survey of deepfake detection methods categorized under the fake image, fake video, and fake audio detection.

A. Fake Image Detection

A new wave of face-swapping technology poses a significant threat to identity and can penetrate systems to get illegitimate access. With many disturbing AI-generated images flooding across various platforms, several detection measures have been proposed in this regard. Most of these fake images are synthesized using high-quality GANs that generate highly realistic content that is very difficult to detect.

Fig.2 shows an example of deepfake image created using a deepfake application called FaceApp. Here U.S President Donald Trump's unhappy face (upper block) has been transformed into a smiling face (lower left) by only using the smile tool in this deepfake application. Next American singer Camila Cabello's face is swapped with that of Donald Trump (lower right) and both these images are challenging to detect as fake.

Fig. 2. Use of Deepfake application FaceApp to generate the fake image

Fake image detection has been treated as a binary classification problem where Convolutional Neural Network (CNN) models have mostly been used for detection [17, 18]. Later adoption of advanced CNN-Xception Network [19] further increased the accuracy of detection. Here they have carried out the extraction of spatial and steganalysis features of the digital content using CNNs. Spatial features involve the identification of visible inconsistencies in the image like facial

blurs, facial texture, artificial smoothness, and contrast difference. Steganalysis help analyze the hidden information in an image through low-level feature extraction. However, with many evolving extensions of GAN, new mechanisms are being adopted. So it is important to improve the generalization ability of detection tools. Instead of focusing on statistical details of a low-level pixel, Xuan et al. [20] proposed the use of Gaussian blur and Gaussian Noise to force classifiers to learn more detailed and meaningful features from improved statistical similarity at the pixel level in images.

Likewise, in [21] a two-phase model is proposed that pairs fake and real images. These pairs then learn from discriminative common fake feature networks (CFFN). The discriminative CFF is used to identify the authenticity of the image. They have highlighted the fact that it is challenging to identify subjects excluded from the training phase using supervised learning. So have introduced a contrastive loss to learn through pairwise learning. The objective of this framework is to detect newly encountered fake images and those generated by a new GAN. Their experimental results have demonstrated to outperform the precision and recall rate of other state-of-the-art methods.

B. Fake Video Detection

When videos are synthesized to swap faces or change facial expressions, the new images do not usually match the lighting conditions or head positioning. This is achieved by geometrically transforming them by rotating, resizing, or otherwise distorting them. This process is guaranteed to leave behind digital artifacts in the resulting image that can make them look doctored. And so various algorithms have been trained to detect these artifacts to check for the authenticity of the content. Li and Lyu [22] propose a method that exposes deepfake videos based on these face-warping artifacts. Instead of self-generating negative data from available datasets like several other models [23,24], which can be a very time consuming and expensive process, they use a simple image processing operation to create negative data. They compare generated face areas with its surrounding using a dedicated CNN model for detection.

Several deepfake algorithms are not able to mimic the normal eye blink rate of a person due to a lack of closed eye images. Wang et al. [25], used contour circle fitting to locate the presence of pupil and thus detect blinks and achieved an accuracy of 96.6%. However, most of these studies were based on datasets created in a laboratory, so their results could not be generalized in a real-world scenario. Li et al. [26], recently proposed the use of CNN along with recurrent neural network (RNN) for detecting eye blinks, to differentiate between a fake and authentic visual. This method outperformed the method proposed in [27], that used landmark detectors to identify the eye corners and eyelid contours for precise detection of eyes and with the use of two different models namely the SVM (Support Vector Machine) and HMM (Hidden Markov Model), detected the rate of eye blinks. This technique was further enhanced to differentiate between a complete and incomplete eye blink. Fogelton and Benesova [28], have designed a model that detects a complete blink, incomplete blink, and no blink, for every frame. This method was implemented on the Researcher's night dataset, and it outperformed all other related work by almost 8%.

Stressing on the fact that several intra-frame inconsistencies and temporal frame inconsistencies can occur across frames in deepfake video generation, Guera and Delp [23] proposed temporal feature analysis of videos to detect fake videos, with the use of a simple convolutional Long-Short - Term- Memory (LSTM) structure. Similar to this, an advanced framework known as SSTNet [29], uses a combination of CNN based spatial feature extraction, steganalysis feature extraction and temporal feature extraction by a single LSTM for fake video detection. This framework achieved a good generalization on the GAN dataset and outperformed previous steganalysis methods when tested on the FaceForensics++ dataset [30]. Fig. 3 shows the working of SSTNet model.

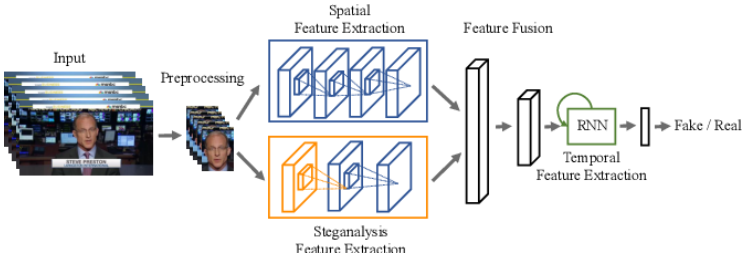


Fig. 3 The proposed framework of SSTNet [29]

C. Fake Audio Detection

Audio manipulation is a combination of AI techniques of deepfakes and general processing like speeding, slowing, cutting, or decontextualizing a video, commonly referred to as Cheapfakes [31]. Resemblyzer is an open-source tool that detects deepfakes by extracting high-level representation of audios that enable developers to compare the two voice samples at any given point of time [32]. Other fake audio detectors focus on the difference in spectrograms - the visual representation of audio, of real and fake audios [33]. However, later an improved CNN based architecture called Wavenet [34] was introduced in various platforms like Text-To-Speech (TTS) and speech recognition. TTS and Voice Conversion (VC) are two types of speech synthesis software. TTS synthesizes human-like speech based on words or phonemes, while VC systems can convert the voice of an utterance into another voice with the same content. DNN models have popularly been used to extract dynamic acoustic

TABLE I
SUMMARY OF PROMINENT DEEPFAKE DETECTION TECHNIQUES

Method	Type	Dataset	Merits	Demerits
CGFace Model[17]	Image	**Real images:** CelebA **Fake images:** generated using PCGAN and BEGAN	automatic extraction of abstract features and use of AdaBoost classifier over softmax	cannot identify hidden features in images
Pairwise learning[21]	Images	**Real images**: CelebA **Fake images:** generated from DCGAN, WGAN, and PGGAN	two phase model is proposed that pairs fake and real images and common fake feature network is trained to distinguish the features between the fake and real images using pairwise learning	may not detect new fake features generated by new GAN
Preprocessing[20]	Image	**Real images:** CelebA-HQ, **Fake images:** generated from DCGAN, WGANGP and PGGAN.	similar image level preprocessing is done to both real and fake images to destroy low level noise cues so model can learn more intrinsic features for classification	not very large increment in performance compared to other similar approaches
Face-warping artifacts[22]	Videos	UADFV, DeepfakeTIMIT	makes use of distinctive artifacts that are obtained during the resolution transformation of images	does not take into consideration temporal inconsistencies across frames
Eye Blinking[26]	Videos	**Real:** 49 real interview and presentation videos **Fake**: generated from above-mentioned videos	Highlighted the fact that deepfake videos have inconsistencies in the human blinking rate compared to real videos	only considers lack of blinking for detection and not the frequency of blinking
SSTNet: Spatial, temporal and steganalysis[29]	Videos/Image	FaceForensics++ GAN-based Deepfakes	extracts low level, mid-level and high-level artifacts of images in videos and also explores the temporal inconsistencies across successive frames	doesn't take into account the inconsistencies in blinking rate
Dynamic acoustic Features[35]	Audio	ASVspoof 2015 database	proposed an HLL scoring method that only makes use of human node outputs for spoof detection	performance is yet to be investigated on replay attacks and spoofing speech produced by WaveNet

features of audiovisuals and label them as real or fake [35].This proposed methodology has shown to outperform the static feature analysis of Gaussian Mixture Model (GMM) classifiers [36]. The summarized version of the most prominent deepfake detection methods has been listed in Table 1.

IV. CONTRIBUTIONS

After an in-depth review of several proposed detection methods for deepfakes, the SSTNet model proves to be the most convincing one due to its flexibility and use of a generalized approach in detecting both the fake images and videos. This model achieved an accuracy level of around 90% to 95%, which is much higher compared to other models trained on the same datasets. This method can further be improvised by incorporating other detection methods like the rate of complete eye blinks and investigating for lip-synching evidence, on datasets released by giant tech companies like Google [37] and Facebook [38] to classify fake videos.

Many researchers have failed to recognize that signature forgery and document forgery is also a matter of great concern. Signatures ensure a great deal of authenticity in various sectors of banking, insurance, healthcare, copyrights and governmental regulatory compliance. Forgery of any form could potentially lead to more immense disastrous consequences. Thus, it is important for researchers to equally invest in this sector as well.

V. CONCLUSION & FUTURE RESEARCH DIRECTIONS

As the deepfake technology approaches towards generating fake content with considerably improved quality, it will likely become impossible to detect them shortly. It is thus, important to immediately respond to the emerging threat posed by deepfakes with great caution. To be able to distinguish between generated and authentic content, organizations can start developing encrypted digital stamps for authentic digital media.

Believability and accessibility have become the most significant drive in deep fake technology. Stopping deepfakes from spreading across massive networks is a considerable challenge and requires social media platforms to step up. They need to develop tools and extensions that can help deal with deepfake content moderation, detection, and prevent their mainstream media coverage. Also confusing deepfakes to generate more flawed output can help detect them easily. This can be achieved with the addition of special noise to digital photos that are uploaded on social media, such that they create a decoy suggesting there is a face when there is none, in reality [39].

To combat deep fake just developing and deploying one or two successful tools is not enough. It will require a constant reinvention of these tools as this technology is evolving at a much faster rate and machine learning plays a crucial role in achieving this. Therefore, the research community should continue their research in developing

countermeasures using machine learning and deep learning to combat the weaponization of deepfakes.

REFERENCES

[1] R. Kumar, J. Sotelo, K. Kumar, A. de Brebisson, and Y. Bengio, "ObamaNet: Photo-realistic lip-sync from text," pp. 1–4, 2017, [Online]. Available: http://arxiv.org/abs/1801.01442.

[2] D. Yadav and S. Salmani, "Deepfake: A survey on facial forgery technique using generative adversarial network," *2019 Int. Conf. Intell. Comput. Control Syst. ICCS 2019*, pp. 852–857, 2019, doi: 10.1109/ICCS45141.2019.9065881.

[3] "Chinese deepfake app Zao sparks privacy row after going viral | Privacy | The Guardian." https://www.theguardian.com/technology/2019/sep/02/chinese-face-swap-app-zao-triggers-privacy-fears-viral (accessed Aug. 26, 2020).

[4] "AI deepfake app DeepNude transformed photos of women into nudes - Vox." https://www.vox.com/2019/6/27/18761639/ai-deepfake-deepnude-app-nude-women-porn (accessed Aug. 26, 2020).

[5] "Deepfakes and the New Disinformation War | Foreign Affairs." https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war (accessed Aug. 30, 2020).

[6] "The Newest AI-Enabled Weapon: 'Deep-Faking' Photos of the Earth - Defense One." https://www.defenseone.com/technology/2019/03/next-phase-ai-deep-faking-whole-world-and-china-ahead/155944/ (accessed Aug. 24, 2020).

[7] "Deep fakes: AI-manipulated media will be 'WEAPONISED' to trick military | Science | News | Express.co.uk." https://www.express.co.uk/news/science/1109783/deep-fakes-ai-artificial-intelligence-photos-video-weaponised-china (accessed Aug. 24, 2020).

[8] R. Chesney and D. K. Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security," *SSRN Electron. J.*, pp. 1753–1820, 2018, doi: 10.2139/ssrn.3213954.

[9] "Don't believe your eyes: Exploring the positives and negatives of deepfakes - AI News." https://artificialintelligence-news.com/2019/08/05/dont-believe-your-eyes-exploring-the-positives-and-negatives-of-deepfakes/ (accessed Aug. 25, 2020).

[10] J. Kietzmann, L. W. Lee, I. P. McCarthy, and T. C. Kietzmann, "Deepfakes: Trick or treat?," *Bus. Horiz.*, vol. 63, no. 2, pp. 135–146, 2020, doi: 10.1016/j.bushor.2019.11.006.

[11] "Understanding Latent Space in Machine Learning | by Ekin Tiu | Towards Data Science." https://towardsdatascience.com/understanding-latent-space-in-machine-learning-de5a7c687d8d (accessed Aug. 28, 2020).

[12] T. T. Nguyen, C. M. Nguyen, D. T. Nguyen, D. T. Nguyen, and S. Nahavandi, "Deep Learning for Deepfakes Creation and Detection: A Survey," pp. 1–12, 2019, [Online]. Available: http://arxiv.org/abs/1909.11573.

[13] I. J. Goodfellow *et al.*, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 3, no. January, pp. 2672–2680, 2014.

[14] "Generative Adversarial Networks: The Tech Behind DeepFake and FaceApp." https://interestingengineering.com/generative-adversarial-networks-the-tech-behind-deepfake-and-faceapp (accessed Aug. 27, 2020).

[15] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," pp. 1–7, 2014, [Online]. Available: http://arxiv.org/abs/1411.1784.

[16] "What is the difference between Generative Adversarial Networks and Autoencoders? - Quora." https://www.quora.com/What-is-the-difference-between-Generative-Adversarial-Networks-and-Autoencoders (accessed Aug. 29, 2020).

[17] L. M. Dang, S. I. Hassan, S. Im, J. Lee, S. Lee, and H. Moon, "applied sciences Identification Using Convolutional Neural Network," 2018, doi: 10.3390/app8122610.

[18] Y. Aslam and N. Santhi, "A Review of Deep Learning Approaches for Image Analysis," *Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019*, no. Icssit, pp. 709–714, 2019, doi: 10.1109/ICSSIT46314.2019.8987922.

[19] C. Google, "Xception : Deep Learning with Depthwise Separable Convolutions," pp. 1251–1258, 2014.

[20] X. Xuan, B. Peng, W. Wang, and J. Dong, "On the Generalization of GAN Image Forensics," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11818 LNCS, no. 61502496, pp. 134–141, 2019, doi: 10.1007/978-3-030-31456-9_15.

[21] C. Hsu, Y. Zhuang, and C. Lee, "applied sciences Deep Fake Image Detection Based on Pairwise Learning," 2020, doi: 10.3390/app10010370.

[22] Y. Li and S. Lyu, "Exposing DeepFake Videos By Detecting Face Warping Artifacts."

[23] G. David and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks."

[24] D. Afchar and V. Nozick, "MesoNet : a Compact Facial Video Forgery Detection Network."

[25] M. Wang, L. Guo, and W. Chen, "Blink detection using Adaboost and contour circle for fatigue recognition," *Comput. Electr. Eng.*, vol. 0, pp. 1–11, 2016, doi: 10.1016/j.compeleceng.2016.09.008.

[26] Y. Li, M. Chang, and S. Lyu, "In Ictu Oculi : Exposing AI Generated Fake Face Videos by Detecting Eye Blinking."

[27] M. Perception and C. Technical, "Eye Blink Detection Using Facial Landmarks," 2016.

[28] A. Fogelton and W. Benesova, "Eye blink completeness detection," *Comput. Vis. Image Underst.*, vol. 176–177, pp. 78–85, 2018, doi: 10.1016/j.cviu.2018.09.006.

[29] "SSTNET : DETECTING MANIPULATED FACES THROUGH SPATIAL , STEGANALYSIS AND TEMPORAL FEATURES YuTao Gao Yu Xiao Alibaba Group China," pp. 2952–2956, 2020.

[30] R. Andreas, D. Cozzolino, L. Verdoliva, J. Thies, M. Nießner, and C. Riess, "FaceForensics ++ : Learning to Detect Manipulated Facial Images."

[31] B. Paris and J. Donovan, "Deepfakes and Cheap Fakes," *Data Soc.*, p. 47, 2019, [Online]. Available: https://site.ieee.org/sagroups-7011/%0Ahttps://site.ieee.org/sagroups-7011/blog/%0Ahttps://datasociety.net/library/deepfakes-and-cheap-fakes/.

[32] "Resemble AI launches voice synthesis platform and deepfake detection tool | VentureBeat." https://venturebeat.com/2019/12/17/resemble-ai-launches-voice-synthesis-platform-and-deepfake-detection-tool/ (accessed Aug. 30, 2020).

[33] "Detecting Audio Deepfakes With AI | by Dessa | Dessa News | Medium." https://medium.com/dessa-news/detecting-audio-deepfakes-f2edfd8e2b35 (accessed Aug. 30, 2020).

[34] "WaveNet: A generative model for raw audio | DeepMind." https://deepmind.com/blog/article/wavenet-generative-model-raw-audio (accessed Aug. 30, 2020).

[35] H. Yu, Z. H. Tan, Z. Ma, R. Martin, and J. Guo, "Spoofing Detection in Automatic Speaker Verification Systems Using DNN Classifiers and Dynamic Acoustic Features," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 29, no. 10, pp. 4633–4644, 2018, doi: 10.1109/TNNLS.2017.2771947.

[36] D. Reynolds, "Gaussian Mixture Models," *Encycl. Biometrics*, no. 2, pp. 659–663, 2009, doi: 10.1007/978-0-387-73003-5_196.

[37] "Google has released a giant database of deepfakes to help fight deepfakes | MIT Technology Review." https://www.technologyreview.com/2019/09/25/132884/google-has-released-a-giant-database-of-deepfakes-to-help-fight-deepfakes/ (accessed Aug. 30, 2020).

[38] "Deepfake Detection Challenge Dataset." https://ai.facebook.com/datasets/dfdc/ (accessed Aug. 30, 2020).

[39] "Scientists Are Taking the Fight Against Deepfakes to Another Level | Discover Magazine." https://www.discovermagazine.com/technology/scientists-are-taking-the-fight-against-deepfakes-to-another-level (accessed Aug. 30, 2020).

Volume 8 Issue 3 (2020) 230-238 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

A study on Rocker-Bogie Suspension for a Planetary Rover Prototype

Anmol Singh, PK Jain
Department of Mechanical Engineering, Delhi Technological University, Delhi-42, India

Article Info
Article history:
Received 15 February 2020
Received in revised form
29 May 2020
Accepted 8 June 2020
Available online 15 June 2020

Keywords: Rover, planetary, Rocker-bogie, Suspension.

Abstract: In light of the notable rocker-bogie mechanism, this paper shows an ideal plan of a rocker-bogie suspension framework so as to ensure high mobile steadiness as well as excellent versatility of a prototype rover vehicle while traversing through rough terrains. It is essentially a suspension arrangement utilized in mechanical automated vehicles utilized explicitly for space investigation. The rocker-bogie suspension-based rovers have been effectively presented for the Mars Pathfinder and Mars Exploration Rover (MER) and Mars Science Laboratory (MSL) missions led by zenith space investigation laboratories all through the world. The proposed suspension framework is presently the most supported structure for each space investigation organization. It is basically a mechanism which comprises of two arms with wheel mounted to each and the two arms are associated through a versatile joint. The current development in design has been studied as well as a different approach towards designing the basic structure of suspension has been done. It has been further verified using various static and dynamic load calculations, solid modelling computer aided design software and simulation software for analysis and testing.

1. Introduction

Suspension is isolated into non-autonomous suspension and free suspension. These two sorts of suspensions are broadly utilized in general vehicles. General vehicles can go on urban streets and expressways but they can't go on unpleasant streets or incredibly uneven streets. The rocker-bogie suspension is proposed to take care of that issue. It is generally used in planetary rovers. Rocker-bogie suspension deals with the locomotion and stability of the rover vehicle. The rover is designed such that it can traverse various difficult terrains and overcome large obstacles in its path while maintaining proper durability and avoiding any unnecessary tumbling in any direction. Most of the components have been made out of aluminum to reduce the overall weight.

The rocker-bogie configuration has no springs or stub axles for each wheel, enabling the rover to move over obstructions, (for example, rocks) that are up to double the wheel's width in size while keeping every one of the six wheels on the ground. Likewise, with any suspension framework, the tilt strength is restricted by the stature of the focal point of gravity. Frameworks utilizing springs will in general tip all the more effectively as the stacked side yields. The framework is intended to be utilized at moderate speed of around 10 centimeters for every second (3.9 in/s) in order to limit dynamic stuns and significant harm to the vehicle while surmounting sizable hindrances.

Rocker, some portion of the term originates from the rocking part of the bigger, forward leg on each side of the suspension framework. These rockers are associated with one another and the vehicle undercarriage through a differential. Comparative with the body, when one rocker goes up, the different goes down. The suspension keeps up the normal pitch edge of the two rockers. One end of a rocker is fitted to a drive wheel, and the opposite end is pivoted to the bogie. The "bogie" portion of the term alludes to the smaller, rearward leg that turns to the rocker in the center and which has a drive wheel at each end. Bogies were generally utilized as burden wheels in the tracks of armed force tanks as idlers circulating the heap over the territory, and were additionally ordinarily utilized in trailers of semi-trailer trucks. The two tanks and semi-trailers presently lean toward trailing arm suspensions.

1.2. Mechanism

So as to go over a vertical hindrance face, the front wheels are constrained against the obstruction by the center and back wheels. The revolution of the front wheel at that point lifts the front of the vehicle up and over the object. The center wheel is then squeezed against the object by the back wheels against the obstruction by the front until it is lifted up and over. At long last, the back wheel is pulled over the object by the front two wheels. During each wheel's traversal of the object, forward advancement of the vehicle is eased back or totally stopped. This isn't an issue for the operational rates at which these vehicles have been worked to date.

Corresponding Author,
E-mail address:17anmol.singh@gmail.com
All rights reserved: http://www.ijari.org

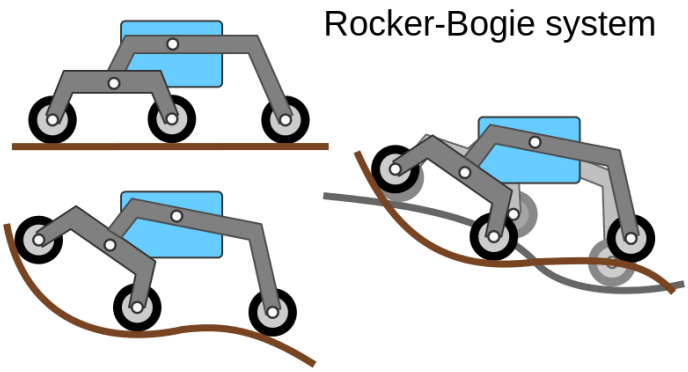


Fig.1: Mechanism of rocker-bogie suspension

Feature	Advantage
No use of springs in suspension	- Doesn't tend to tip/tilt - Can climb high obstacles - Rover doesn't oscillate too much
Independent left/right traversing	- Both the sides can independently traverse different obstacles - Normal reaction remains in control
6 wheels, instead of 4	- Increases normal reaction, traction and stability - Increases independent behaviour of rover - Keeps the rover powered in case of failure in wheel/motor
Differential	- Reduces the main body tilting by half - Jerk experienced by one side is transferred to another in the opposite direction, thus increasing normal reaction

2. Literature Review

2.1 Introduction

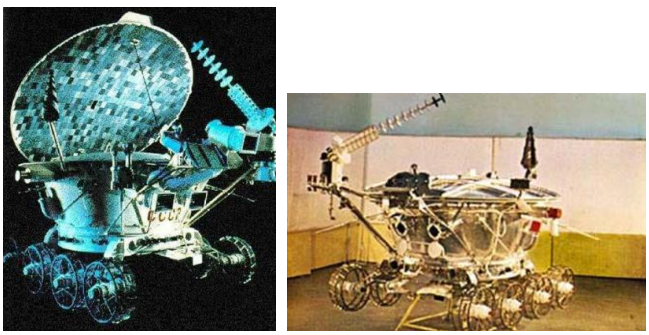
Inception of rocker bogie suspension framework can be followed to the improvement of planetary rover which are portable robots, particularly intended to proceed onward a planet surface. Earlier, movers were tele operated like the Lunokhod I while late ones are completely self-sufficient, for example, FIDO, Discovery and as of late created Curiosity Mars exploration vehicle. The rovers should have been extremely powerful and dependable, as it needs to withstand dust, solid breezes, corrosion and huge temperature changes under baffling conditions. Most extreme rover vehicles stay controlled by batteries which are energized by solar photovoltaic boards during the day introduced over their surface. The motion arrangement of rover vehicles stays urgent to empower it to arrive at

target destinations, direct research, and gather information and to situate itself as indicated by the interest. There are three primary sorts of rover vehicle locomotion invented so far for example wheeled, legged and caterpillar motion. The primary contrast between the different structures of planetary robots lies in the kind of movement framework. Considerably subsequent to creating numerous legged and crossover robots, most analysts still spotlight on wheeled design for rover vehicles in view of its locomotion ease and focal points and among wheeled movement plan, the rocker bogie suspension framework-based structure stay generally supported. The antiquated FIDO rover and the Sojourner contain 6 autonomously directed and driven wheels suspended from a rocker-bogie system for most extreme suspension and ground clearance. Rocky Seven Rover has a comparative suspension framework simply contrast in front wheels. The Nanorover and Nomad Rovers have four guided wheels suspended from two bogies and CRAB Rover uses two parallel bogie-systems on each side to conquer objects and huge gaps. To the extent the underlying examination is concerned, the product advancement looks for an ideal in the restricted solution space given an underlying arrangement and infer a scientific model to sum up rover suspension parameters which characterize the geometry of the rocker bogie framework. The target behind advancement of rocker bogie suspension framework is to build up a framework which limits the vitality utilization, the vertical removal of the rover vehicle's center of gravity and its pitch angle. In this exploration our undertaking is to move the significant favorable circumstances installed with the rocker bogie framework into customary vehicles so as to expel solace and complexities present in regular suspension framework as a rule and suspension arrangement of substantial vehicles specifically.

2.2 Development in the design of Rocker-Bogie for Planetary Rovers

2.2.1 The Rocker-Bogie suspension for Planetary Rover designed by Soviets

In 1970, Luna 17 carried Lunokhod 1 which is the world's first lunar exploration vehicle onto the lunar surface [7]. As shown in Figure 1 (a), Lunokhod 1 is about 756 kg and the body size is 4.42m × 2.15m × 1.92m. The structure of this vehicle is divided into upper and lower parts. The upper part is the sealed instrument compartment manufactured from magnesium alloy and the lower part is the automatic walking chassis which the eight wheels are independently drove. It can climb the 30-degree slope. It can also cross the 40-cm barrier and the 60-cm gully. The ground moving speed is 1km/h or 2 km/ h.



(a) Lunokhod 1 (b) Lunokhod 2

Fig.2: Two kinds of Lunokhod

The Lunokhod 2 was carried by the Luna 21 in 1973 and landed in the eastern part of the Moon [8-9]. As shown in Figure 2 (b), the configuration of the vehicle is similar to Lunokhod 1. However, the structure was improved, the speed of movement was increased, and the scope of activities was bigger than the Lunokhod 1.

2.2.2 The Rocker-Bogie Suspension for Planetary Rover Designed by Swiss

In 2002, Swiss Federal Institute of Technology and Automation Systems Laboratory (ASL) developed a prototype of six-wheel detection vehicle called shrimp [10]. The shrimp is about 3.1kg, which the front fork-parallelogram rocker suspension is used. As shown in Figure 2, the shrimp was climbing stairs. The stability of the shrimp is good and the traffic ability of the shrimp is strong. But the area of carrying in this vehicle is greatly limited due to the limitation of the suspension structure. The suspension structure is also complicated.

Fig 3 (a): The Shrimp

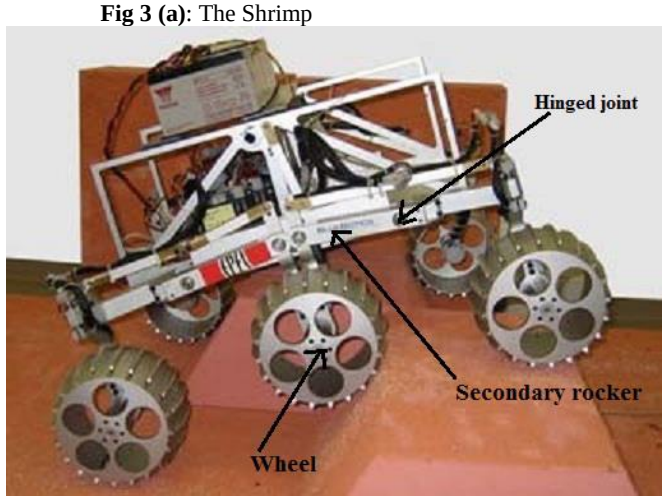


Fig. 3 (b): The six rounds of the prototype vehicle

In 2004, Switzerland ASL developed the six rounds of the prototype vehicle called crab, which the bilateral double parallelogram rocker suspension is used [11]. Each side of the suspension is made up of two sets of the parallelogram secondary rocker hinged at the middle wheel and a main rocker. The pole of each set of the secondary rocker is connected with the wheel. The middle pole of secondary rocker is fixedly connected with the main rocker. The main rocker is connected to the vehicle body with a differential mechanism. As shown in Figure 3, the crab was crossing the vertical wall. The stability of this vehicle is good and the terrain adaptability is strong. The platform area of carrying is much better than the shrimp. But the suspension structure is still complex.

2.2.3 The Rocker-Bogie Suspension for Planetary Rover Designed by American

In 1987, JPL of NASA presented Rocky1 that is a basic prototype of a small 6-wheel rover which firstly used the rocker-bogie suspension [12]. The wheel radius is 4 cm. The wheel to ground distance at the same side is identical. All the wheels can be independently driven. The portion between the rear wheel and hinged joint is the primary rocker. The portion between front wheel and hinged joint is the secondary rocker. The height of free hinged joint is higher than the axis of the wheel. As it was showed in Figure 4, the rover was traveling through an obstacle without any problem. This experience revealed that the negotiation ability and the traffic ability of this rover is strong. The revolution angle of main rocker cannot be very large due to the limitation of suspension system.

Later, JPL developed lots of rovers. The Rocky 8 is the newest rover [13]. The weight reached up 30 kg (Showed in Figure 5). The wheel to ground distance at the same side is also identical. Each wheel can be also independently driven. The height of hinged joint is also

greater than the centre line of the wheel. The Rocky 8 can carry more valid loading weight. It can travel at a long distance and avid obstacles.

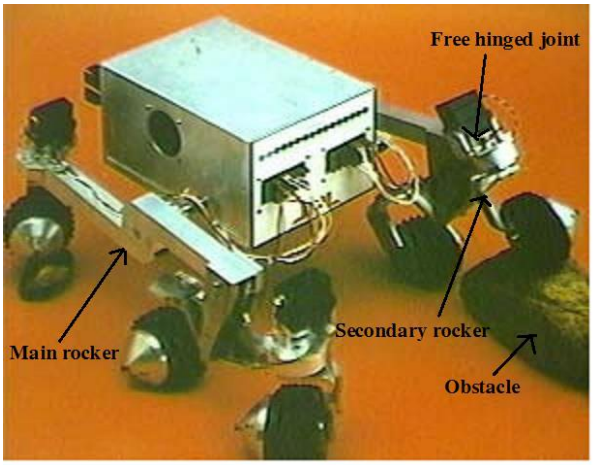


Fig.4: Rocky 1

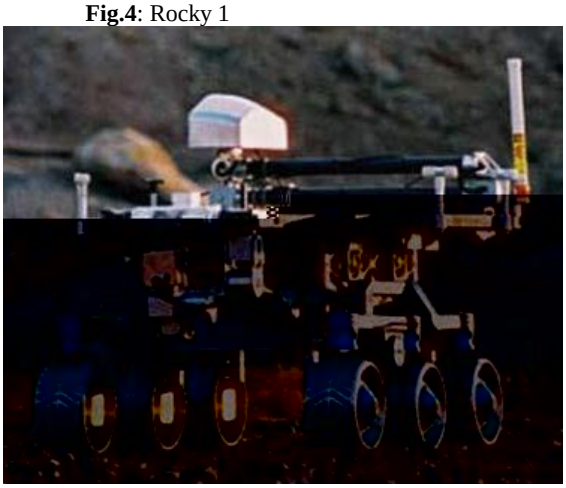


Fig.5: Rocky 8

Future combat system (FCS) designed by the U.S. Army combines many kinds of systems, which is a multifunction, networking, light and intelligent weapon system [14]. This system is gradually equipped by the U.S. Army, as the technology matures. The robot vehicle is called mule as the main transport equipment of FCS, which is used to transport weapons, ammunitions and supplies for army's [15]. One mule can support two infantry classes and also can accompany them on the complex terrain. The primary mission of mules is transporting equipment and rucksacks about 900~1100 kg and transporting the wounded soldier to safe place.

Fig 6: Mule

From the Figure 6, it showed that the mule could travel on the complex terrain with high manoeuvrability. "Mule" was equipped with independent articulated suspension of 6×6 chassis. Mule used brake steering device and hybrid electric drive system. The hub of each tire is equipped with an electric motor. Mule has a high resistance to destruction. If a tire is damaged or destroyed, the remaining five tires still be able to maintain sufficient mobility. The height of obstacle which the mule can across reached 1.5 m, which is much more than any other active wheeled or tracked vehicle equipped by the U.S. Army. The width of the trench which the mule can across reached 1.5 m. The slope angle is more than 40 degrees. The wade depth is at least 1.25 m.

2.2.4 The Rocker-Bogie Suspension for the Planetary Rover Designed by Chinese

National Defense University of Science and Technology designed a new type of folded-deployed suspension, known as the two-crank-slider suspension [16]. As shown in Figure 7, the left and right suspension is symmetrical and each suspension is composed of the main rocker, two cranks, two connecting rods, two slider and torsion bar spring. The left and right suspension is connected with differential mechanism through the spline. The center of the differential mechanism is attached to the body. Both side suspensions are rotatable relative to the vehicle body. Each wheel is driven by a wheel pump. But the crank rotation angle is limited.

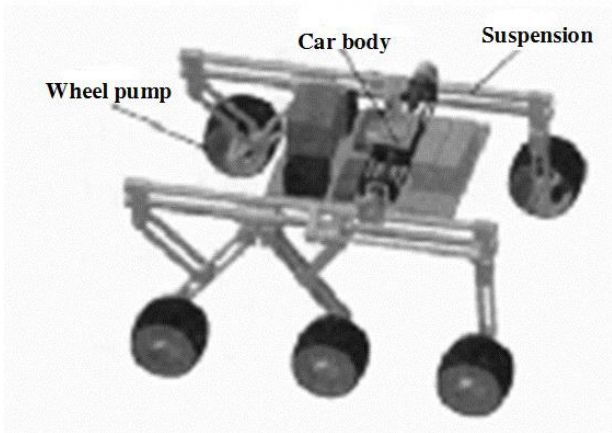


Fig 7: A new type of folded-deployed suspension

Harbin Institute of Technology created the rocker-style inspection vehicles [17]. This suspension is similar to Rocky 1. The supplementary rocker can revolve about the hinge on the primary rocker and the right and left primary rocker arms can realize the differential movement around the differential mechanism so that the 6 wheels always touch the ground at the same time. In addition, the mobile sub-system consists of the 6 full-drive rounds. On the one hand, it increases adhesion and traction. On the other hand, if one or even two wheels fail to work, the remaining wheels continue to run.

2.2.5 The Rocker-Bogie Suspension for Planetary Rover Designed by India

The rover vehicle's mass was around 27 kg (60lb) and was intended to work on solar oriented power. The rover was to proceed onward 6 wheels navigating 500 meters on the lunar surface at the pace of 1 cm for each second, performing nearby examination and sending the information to the Vikram lander, which would have communicated it to the Earth station. The rover was equipped with the following for navigation:

- Stereoscopic camera-based 3D vision: two 1-megapixel, monochromatic NAVCAMs before the rover to give the ground control group a 3D perspective on the encompassing territory, and help in way arranging by producing an advanced elevated model of the landscape. IIT Kanpur added to the improvement of the subsystems for light-based guide creation and movement making arrangements for the rover.
- Control and motor dynamics: the rover design has a rocker-bogie suspension system and has six wheels, each of which are driven by independent brushless DC electric motors. Steering

system is achieved by differential speed of the wheels also called skid steering.

The expected operating time of Pragyan rover was supposed to be one lunar day or around 14 Earth days, as the onboard electronics were not designed to sustain the freezing lunar night. The power system had a solar-powered sleep/wake-up cycle applied, which could have led to longer service time than was planned.

Fig 8: Pragyan Rover

- Dimensions of rover: 0.9 * 0.75 * 0.85 m
- Power required: 50 Watts
- Travelling speed: 1 cm/sec
- Planned mission operating time: ≤14 days

2.3 Future scope

Rocker-bogie suspension system can be implemented in exploration rovers for the following purposes: -

- It can be used in rovers designed for assisting astronauts in space operations and it can act as a path finder.
- It could be useful for implementation in rovers for space missions too as recently it was used in NASA's Mars Rover Curiosity and ISRO's Lunar Rover Pragyan. The following mechanism takes into consideration the roughness of the surface it is driven on.
- The rovers with rocker bogie generally have larger wheels as compared to objects; therefore, it can easily function over maximum of the Martian rocks.
- It can be used in rovers made for excavation purposes in coal mines.
- It can be used in rovers that act as spy robots and for military operations too due to capability of working in off-road conditions.
- It can be used in heavy machinery cranes and bulldozers for ease in traversing through rough terrains

3. Design Methodology and Calculations

3.1 Overview of design methodology

Assumptions:

• We chose a wheelbase of 100cm according to the general limitations in rover competition specific rulebooks.

• It was assumed that all of the load on chassis is concentrated at a point at the center, i.e. also the hinge point of rocker.

• A high ground clearance of 50cm was assumed so that the rover easily encompasses larger obstacles over the wheels.

• The shape of links was designed on AutoCAD to minimize the possibility of collision between bogie and arm during climbing and descending.

• The middle wheel is positioned at the center to distribute load on the wheels equally

• All the rest of the lengths, including the lengths of the links were automatically obtained by restricting the parameters mentioned above on the CAD model.

• Rovers operate on extremely low speeds(3km/h), therefore any dynamic force due to acceleration of the vehicle is neglected

A study of the climbing capacity of different bogie hinge point configurations was done to optimize the value of traction and climbing capacity, as follows: -

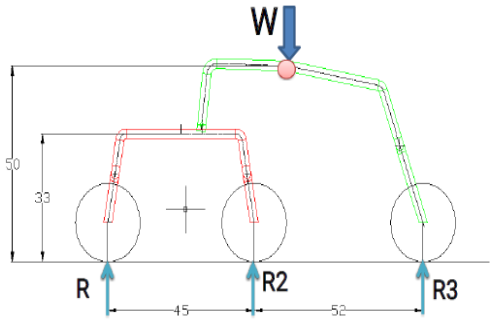


Fig 9: Ground Reactions on horizontal plane

R1, R2, R3 are the ground reactions of wheels on one side of the rocker bogie arrangement

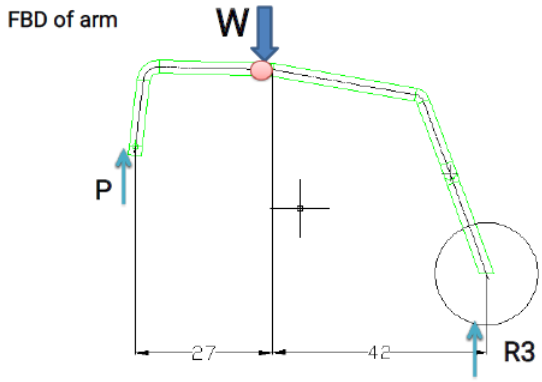


Fig 10: FBD of arm

P is assumed to be a combined effect of R1 and R2

Taking moment about hinged point

⇨ P*27=R3*42

Therefore almost 60% of load W is transferred to point (having pin connection between bogie and arm)

⇨ P=0.6*W

And around 40% of load W is transferred to the contact point of the rear wheel

⇨ R3=0.4*W

This leads to a stable balance of the chassis

Case1

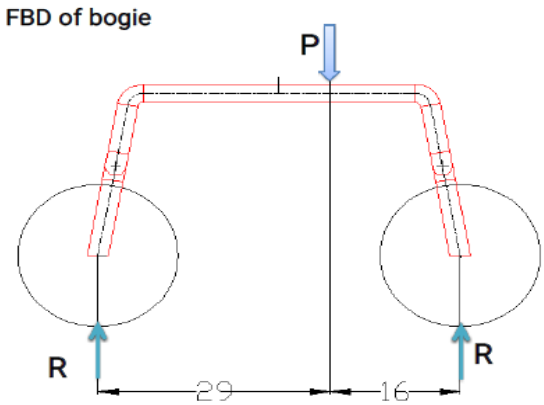


Fig 11: FBD of bogie

In this case the hinge point is set a little offset from the center to analyze the climbing capacity of rover in such configuration

Taking moment about pin connection of bogie and arm

⇨ R1*29=R2*16

Therefore almost 65% of load P is transferred to the middle wheel

⇨ R2=0.65*P

And 35% of load P is transferred to the front wheel

⇨ R1=0.35*P

- R3=0.4*W
- R2=0.6*0.65*W
- R1=0.6*0.35*W

3.1.1 Climbing analysis

3.1.1.1 Position1

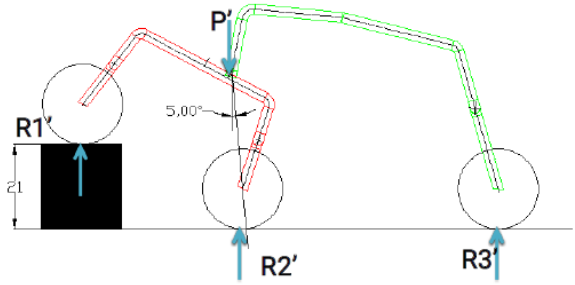


Fig 12: Climbing analysis at position 1

For climbing the obstacle without losing contact of any wheel from ground, it is necessary that the line of action of reaction force R2'should always behind the line of action of force P'. This will generate a counter clockwise direction moment so that the front wheel will always keep in contact with ground.

- In limiting case the line of action of both forces are in a line this mean the reaction force R1' is equal to zero,
- By considering small counter clockwise moment, the maximum climbing step height is limited to 21cm (max).

3.1.1.2 Position 2

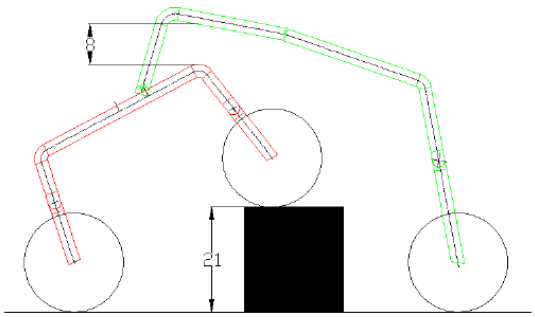


Fig 13: Climbing analysis at position 2

- In step climbing(position2) with 21cm obstacle height, there is no clash between bogie and arm. (8cm clearance)

With this type of configuration (case1)

- Maximum step climbing is obtained as 21cm.
- Better climbing capacity.
- Plenty of clearance between bogie and arm in the Position 2, while declining that means lesser risk of collision.

3.1.1.3 Case 2

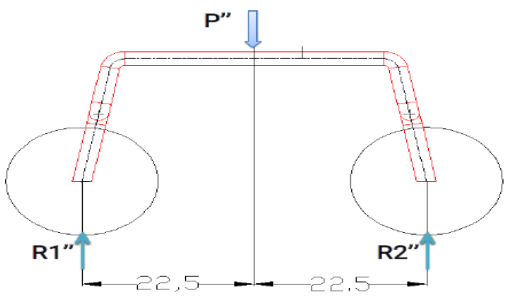


Fig 14: FBD of bogie

In case 2, pin connection of arm with bogie is at center between first and second wheel. And the load P is 60% of the load W.

Therefore R1''=R2''

- R3=0.4*W
- R2=0.3*W
- R1=0.3*W

This configuration provides better traction as the values of normal reactions on wheels are closer to each other.

3.1.2 Climbing Analysis

3.1.2.1 Position 1

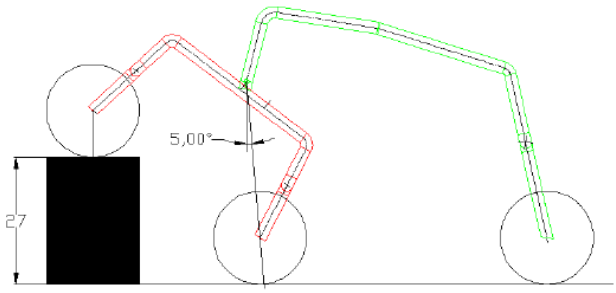


Fig 15: Climbing analysis at position 1

- By considering small counter clockwise moment, the maximum climbing step height is limited to 27cm (max) with this type of configuration.

3.1.2.2 Position 2

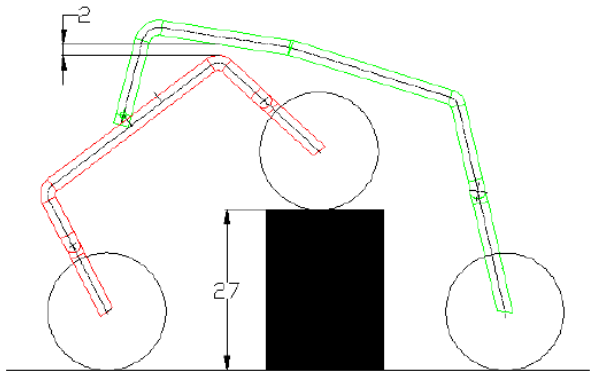


Fig 16: Climbing analysis at position 2

In step climbing (position2) with 27cm obstacle height, there is no clash between bogie and arm (2cm clearance). With this type of configuration-

- Maximum step climbing is obtained as 27cm
- Better traction as compared with case1
- Only 2cm of clearance left in Position 2 while declining that leads to greater risk of collision between bogie and arm as compared to Case1.

3.2 Components in the assembly

Our vehicle's structure has been primarily divided into three systems that enables easy identification and troubleshooting of failures and maintenance by enhancing accessibility to its various constituents. The three parts include: Chassis, Suspension system and Wheels. Solidworks model as well as Finite element analysis for different stress values have been shown for the given components.

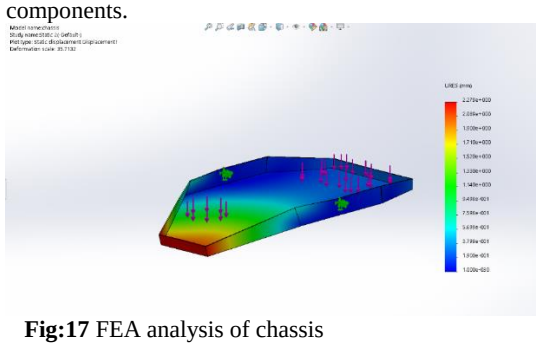


Fig:17 FEA analysis of chassis

3.2.1Chassis

A robust octagonal structure has been chosen from various designs as it provides greater clearance for the movement of driving members and links and also a wider space for placing electrical components and differential mechanism also, walls of this structure provide a safer and dust free environment for electrical components by enveloping them under an acrylic sheet. A differential mechanism has been provided to distribute the various forces while traversing different terrains. The main advantage of this mechanism is that the pitch angle of chassis is the average of the pitch angles of right and left sides of suspension system.

3.2.2 Suspension System

The suspension system of rover consists of rocker bogie mechanism. The rocker-bogie configuration has no springs or stub axles for each wheel, enabling the rover to move over impediments, for example, rocks, that are up to double the wheel's diameter in size while keeping every one of the six wheels on the ground. Rocker Bogie suspension provides the purpose of stabilizing the rover at rough terrains. By virtue of rocker and bogie links, this suspension system provides optimal weight distribution as well as high climbing capacity while smoothly traversing through terrains. The system has been provided with two hinge point configurations that allow us to manually adjust between climbing capacity and magnitude of traction. As with any suspension system, the tilt stability is restricted by the height of the centre of gravity, therefore we have provided with heavy wheels and partially in-wheel drive motors to lower the centre of gravity while maintaining high ground clearance. Also, to attain full traction on wheels, the links have been bent such that most of the rover's weight is concentrated on the wheels.

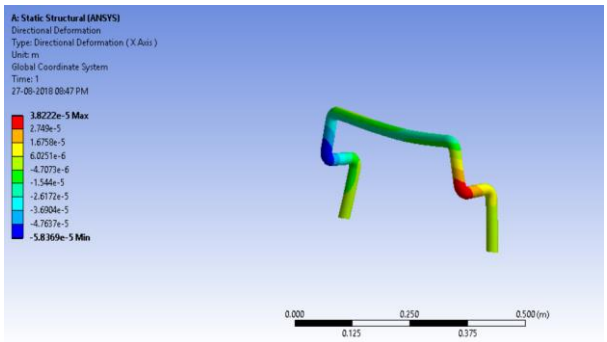


Fig 18: FEA analysis of bogie link

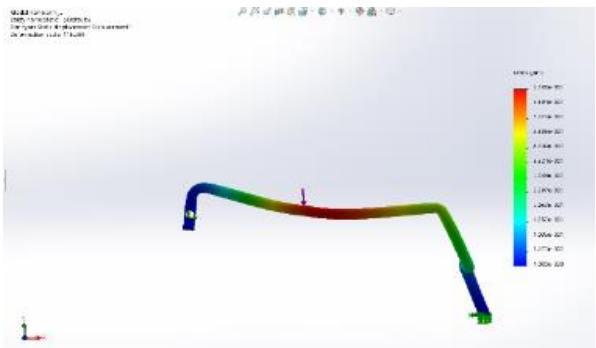


Fig 19: FEA analysis of rocker link

3.2.3 Differential-Purpose

- In a real rover, the rocker bogie mechanisms of left and right sides are connected to each other and to the body through a mechanism called differential.
- The differential is what keeps the body level. Relative to the body, when one rocker goes up, other goes down.
- Relative to the ground, the body pitch angle is always half of the angles of left and right rockers.
- If we attach the rockers to the body with an axle or two pivot points, the body will tip forward or backward until it hits the ground. The differential makes the rover body to remain stable and prevents it from tilting.
- The differential makes it possible for one side or both sides of the rover to overcome keeping the body stable with low pitch angle.

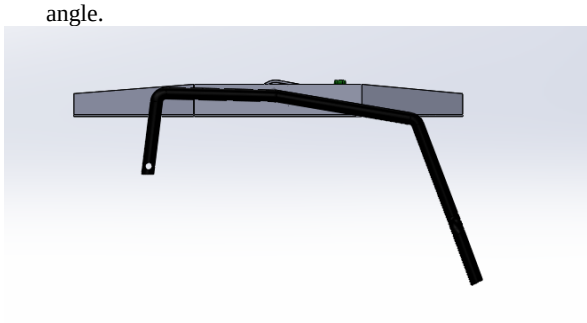


Fig 20 : Balancing of chassis roll angle

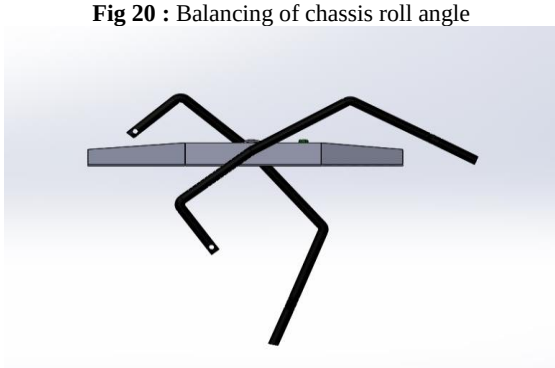


Fig 21 : Balancing of chassis pitch angle

Table 2: Types of differential mechanisms

Bevel Gear Differential	Differential Bar
Bevel Gear assembly housed inside the body.	Assembly consisting of different bars of short length.
Two bevel gears connect to two rocker arms and the third (middle) gear connects to the body.	Differential bar is pivoted on its middle point to the body and its ends are connected to the rocker arms through some short length links.
Easy to manufacture and assemble.	Difficult to fabricate, converting motion is complex.
Reliable and has more strength due to steel alloy bevel gears.	Not much reliable.
The assembly is compact and can be easily housed inside the body.	The assembly needs separate space to work and interferes with other components.

On the basis of the above comparison between bevel gear differential and differential bar, bevel gear differential was chosen because of the advantages in its mechanism. A separate compartment for the differential mechanism inside the rover body for preventing the dust particles from entering. Therefore, the gears can work smoothly.

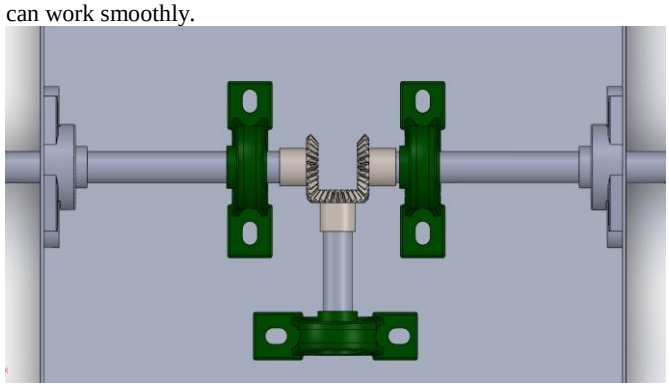


Fig 22: Configuration of differential

Table 3: Configuration of differenetial

Part	Material	Quantity
Bevel Gears	EN8, Hardened	3

Shafts	Aluminium	3
Plummer Blocks	Cast Iron	4-6

3.2.4 Wheels and hub

Wheels have the important function to support the load of the rover and help navigate through different terrains, like sandy, non-cohesive soil as well as hard, dry terrain. In addition to these features, wheels also perform the function of providing an elastic behaviour to increase contact force and flexibility as well as reducing overall vibrations and jerks. Balloon tyres are used. These tyres provide with excellent strength at light weight. However, the main reason for using balloon tyres is the flexibility in traction that they provide. On the tyres we have used rubber strips for better ground control and extra traction. The traction in these types of tyres is controlled by lowering the air pressure inside them.

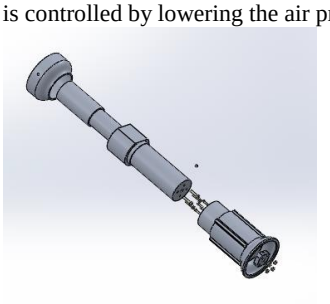


Fig 23: Wheel hub

Fig 24: Balloon tyre

The wheel hub is custom-made from aluminium and is designed to completely hold the motor in itself. In-wheeling helps in the optimum utilization of space while reducing the danger of motor-damage. The intricate hub is manufactured using CNC machining.

I owe a debt of gratitude to our project guide, Mr. PK JAIN, Associate Professor, Department of Mechanical Engineering, DTU, for incorporating in me the idea of a creative Research Project, helping me in undertaking this project and also for being there whenever I needed his assistance.

References

[1] NA Cabrol, G Chongdiaz, JM Dohm, et al. Atacama I: Science Results of the 1997 Nomad Rover Field Test in the Atacama Desert, Chile. Proceedings of the Lun Planet Sci Conf, F, 1998
[2] S Hayati, R Volpe, P Backes, et al. The Rocky 7 rover: a Mars science-craft prototype. Proceedings of the IEEE International Conference on Robotics and Automation, 1997
[3] R Volpe, J Balaram, T Ohm, et al. Rocky 7: a next generation Mars rover prototype [J]. Advanced Robotics, 11(4) 1996, 341-358.
[4] A Ellery. Environment–robot interaction—the basis for mobility in planetary micro-rovers [J]. Robotics & Autonomous Systems, 51(1), 2005, 29-39.
[5] R Siegwart, P Lamon, T Estier, et al. Innovative design for wheeled locomotion in rough terrain [J]. Robotics & Autonomous Systems, 40(2-3), 2002, 151-62.
[6] Research Development of Deep Space Detection Vehicles and Lunar Rover of China [J]. 2009
[7] S Kassel. Lunokhod-1 Soviet Lunar Surface Vehicle [J]. Rand Corporation, 1971
[8] M Lindroos. The Soviet Manned Lunar Program [M]. MIT Open Course Ware. 2011
[9] Ducrocq A. Lunokhod-2. I (Translunar flight plan of Luna 21 probe, describing Lunokhod structure and lunar descent procedure) [J]. Air Cosmos, 1973 44-47.
[10] R Siegwart, P Lamon, T Estier, et al. Innovative design for wheeled locomotion in rough terrain [J]. Robotics and Autonomous systems, 2002 40(2) 151-62.
[11] T Thueer, A Krebs, R Siegwart. Comprehensive locomotion performance evaluation of all-terrain robots; proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, F, 2006 [C]. IEEE.
[12] DB Bickler. The new family of JPL planetary surface vehicles [J]. Missions Technologies & Design of Planetary Mobile Vehicles, 1993 301-306.
[13] PG Backes, JS Norris, MW Powell, et al. Multi-mission activity planning for Mars lander and rover missions [J]. 2004 2(2)877-886.
[14] R Dietterle. The future combat systems (FCS) overview[C]. Proceedings of the Military Communications Conference, Milcom, F, 2005
[15] DB Byers. Multifunctional Utility/Logistics and Equipment (MULE) Vehicle Will Improve Soldier Mobility, Survivability and Lethality [J]. Army Al & T, 2008
[16] J Shang, Z Luo, X Zhang. Design and prototype development of a lunar rover with two-crank-slider suspension [J]. Zhongguo Jixie Gongcheng/China Mechanical Engineering, 18(3), 2007, 348-51.
[17] JG Tao, J Li, B Li. Design of single drive folded-deployed suspension for six-wheeled rocker-bogie explore rover [J]. 2012, 21-26.
[18] Ramesh, Sandhya. Why Chandrayaan-2 is ISRO's 'most complex mission' so far. Print ISRO to send first Indian into Space by 2022. Department of Space 6, 2019.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

A survey on enhanced RPL: Addressing the mobility in RPL

Rajesh Kumar Yadav
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rkyadav@dtu.ac.in

Abstract-The Internet Of Things(IoT) is emerging as a necessary tool for the modernization of world. With the advancement in IoT comes the new challenges in many areas. One area among them is routing. Though routing in IoT is challenging but the more challenging task is routing in low power lossy networks (RPL). There are some critical applications for example healthcare, which requires the mobility issue to be resolved. Mobility in RPL is a new research area for many scholars and number of protocols are developed for the same. Some work integrate mobility in a way that avoid any disconnection of moving node before finding new connection while other just allow fast recovery after disconnection. This work is analysis of various mobility associated RPL protocols. The main focus of this paper is on the energy consumption, handover delay, signaling cost and route stability of various algorithms. The presented work also focuses on whether the protocol is reactive or proactive. This work gives a detailed conclusion on development of mobility in RPL and categorized the studied protocols on the basis of their property. To some extent protocols studied are also compared to each other and tells which one is better.

Keywords— Routing protocol for low power lossy network, Mobility, Handover delay.

I INTRODUCTION

The internet has become an integral part of our life. With the development of smart appliances comes in the technology IoT(Internet of Things)[1] which is expanding in every area possible. With development and enhancement in IoT many global organizations has worked to provide guidelines, rules and specification to create advanced IoT applications that can improve our life. A vast range of applications are now part of IoT. Most of the devices in IoT are low battery, mobile objects that have limited resources. To achieve desired result, a well integration of mobility and resource utilization is required. Wireless sensor networks [2] is one among those areas. Wireless sensor network is considered as a set of resource constrained devices which are dispersed in a confined physical area and collects information about surroundings. All the node in WSN send the information to a central repository where analysis is performed. Wireless sensor networks has played a vital role in many of the applications like healthcare, traffic control, military etc. Since all these applications require motion of resource constrained devices, it give researchers a new area to work upon. The network type of WSN is a resource constrained called Low power and Lossy Networks(LLN)[3] and routing and resource utilization in such kind of network has always been challenging.

Nimisha Awasthi
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
awasthi.nimisha97@gmail.com

In these types of networks all the equipment like router and interconnection are resource constrained. These types of networks are characterized by low throughput, low processing speed, high data loss and limited memory. The end devices in these types of networks mainly collect data and send it to a central repositories for analysis. Since it requires sending of data from end devices to central repositories, appropriate routing methodology need to be developed to maximize the resource utilization and avoid data loss.

Mobility[4] in any network provide many advantages by introducing more innovative services. However, mobility brings in some new challenging issues[5] that need an appropriate approach to resolve. Initially the RPL protocol was developed for low power device which do not take into consideration the mobility. It uses the strategy of self-healing[6] which raise issue like data loss, increased handover delay, route instability, frequent change in network topology and increased flow of control messages. To overcome these shortcoming protocols are developed to integrate mobility with efficiency. With these developments the various important parameters like handover delay, energy consumption, route stability are improved. Most of the mobility-integrated protocol can be classified either as reactive protocol or proactive protocol. In proactive protocols a new parent node is found before a mobile node gets disconnected from its parent while in reactive protocol a parent node is found after disconnection but the handover delay is reduced.

The proposed work is presented in the following layout: Section (II) gives an overview of RPL protocol and RSSI value. Section (III) presents various enhancement made to RPL for mobility. Comparison and Classification is presented in Section (IV) and Conclusion is explained in Section (V).

II RPL AND RSSI OVERVIEW

A). RPL Overview

Routing in Low power Lossy networks (RPL) as the name suggests is mainly for LLN networks. It is a distance vector routing protocol which is mainly concerned with sending of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

data from end devices to central repositories that forms a tree-based Destination Oriented Directed Acyclic Graphs(DODAG) and aims to establish a seamless network. The RPL consists of control messages; DIO, DAO, and DIS[19] which are used for the construction of DODAG tree.

- DAG Information Object(DIO): DIO are multicast messages sent for DODAG discovery and maintenance. Each Nodes on expiry of trickle timer transmit a DIO messages periodically to inform the current state of the network. The initial construction of DODAG tree is done using DIO message

- Destination Advertisements Object (DAO): DAO message is used to transfer traffic from root to downward in non-storing mode. In this case, if a node is having some data to be sent to some other non-root node then data is first transferred to root node and then root node transfer the data to destination node. DAO message is send by a node to its parent node.

- DAG Information Solicitation (DIS): DIS message is designed to be sent by a new node to join the DODAG.

The DODAG building process is started by root node or sink node. The root node sends DIO message to all the neighboring nodes. The DIO message contains Objective Function (OF) [7] and other details. Each neighboring node on receiving DIO message decides whether it want to join DODAG or not based on OF. A DIO message can be received from more than one node, the parent node selection is done based on node rank. The new nodes, which joins the DODAG, construct new DIO message by calculating their node rank and send it to their neighbor node. In this way the whole DODAG tree is formed.

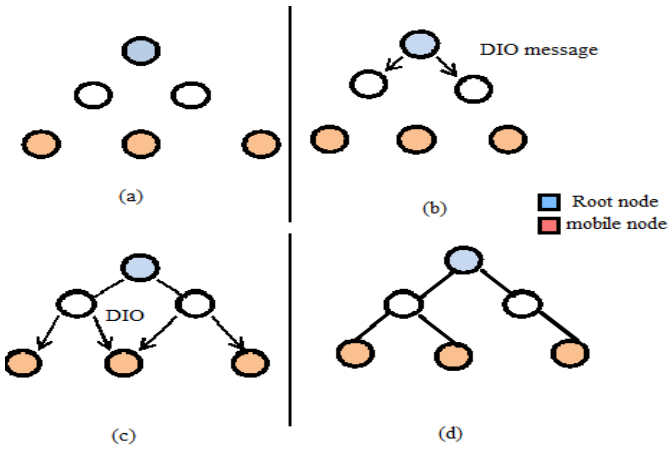


Figure 1 DODAG Construction of RPL tree.

When a new node wants to join the DODAG it sends a DIS message to all the neighboring nodes. In response, the new node receives DIO message and according to the parameters, it decides which node to select as parent.

The RPL DODAG tree construction is shown in fig 1. As shown in the figure 1(b), first the root node sends the DIO message to all the neighbor nodes and so on untill the tree is constructed. In the fig DIO message to only those node is shown whose parent is not selected but in actual DIO message is sent to all the neighboring node. Since RPL is for low power networks, it must ensure that energy consumption is low. In order to achieve so RPL uses trickle timer [8] to slow down the control message.

RPL selects routing path on the basis of various factors like resources available, data loss rates, battery life and other factors and hence ideal for LLN networks. The objective function of RPL is designed taking into consideration of above factors.

B). RSSI Overview

Received Signal Strength Indicator (RSSI) value is calculated by the parent node using the messages received from child node. It gives the strength of signal received and is measured in decibel. The range of this value varies from 0 dB to -100 dB. The closer the value to 0dB the better the signal strength. Generally a RSSI value of -60 dB is considered to be a good signal strength. The more the value is negative the weaker the signal strength and higher the data loss. Most of the mobility integrated RPL protocol uses this value to get reliable paths.

III MOBILITY SUPPORT IN RPL

RPL was originally designed for low power lossy networks and it was assumed that network topology change would be minimal, but due to mobility an inconsistency arise in topology. When this happens, the mobile node becomes detached from it's all the connected nodes (parents and children). In this scenario, RPL react with self-healing strategy by locally repairing the path by receipt of DIO message after the trickle time. In this way in order to avoid data loss, RPL tries to find an alternative path which is not optimal.

In Mobility Enhanced RPL (ME-RPL) [9], unlike traditional RPL firstly a distinction is made between static nodes and mobile nodes. Mobility is included in the control messages using flag to increase route stability and avoid unnecessary disconnection of routing paths. Mobile node can be only leaf node and never a parent node. Further to improve recovery from the disconnection due to mobility, trickle timer period is kept dynamic. It depends on the mobility frequency of the network. The higher the mobility in network the less the trickle timer period and accordingly DIO messages are triggered. Sometimes due to high mobility many DIO messages are triggered which may cause high energy consumption of the node. It is a reactive protocol as it finds new parent node after the child node has lost its connection. Many other parameters like handover delay[10] which is very important for many critical applications, route stability and others are also not considered.

In [11] the authors enhances and extends the MERPL protocol. EMA-RPL starts with the same condition as of [9] that a node which is moving will be distinguished from a

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

static node using mobility flag but the approach used in this is very different from that of [9] as it provides a proactive protocol. Before the leaf child node gets disconnected from its parent node due to motion, the parent node gets to know about this mobility using RSSI [12] value. This value is calculated from the messages received from leaf child node and if this value is not above a threshold value then parent node is alerted and before disconnection, it starts finding new parent node. The entire process is carried out by the parent node itself that is from discovering a new parent after detecting mobility to notifying about the new parent to mobile node, hence it saves energy consumption of mobile node. In terms of handover delay, energy consumption and other parameters it beats both RPL and ME-RPL but the routes generated using this protocol are highly unstable. There is low route stability as the new parent node is selected when the node is still in motion.

An Extended KP is used in [13] which gives routes that are more stable as compared to [11]. The approach is similar to EMA-RPL up to some extent as mobility detection is same as that of EMA-RPL and beside this it is also a proactive approach, differentiating between mobile nodes and static nodes. The difference is that the new parent node is found by the mobile node itself using its co-ordinates and direction of motion. The co-ordinates are enhanced using Extended Kalman Filter. The route stability as compared to EMA-RPL is high and disconnection rate is reduced. The thing of concern is that most of the calculation and updation are performed by mobile node itself which increase mobile node energy dissipation. Figure 2 gives energy consumption of above explained protocols.

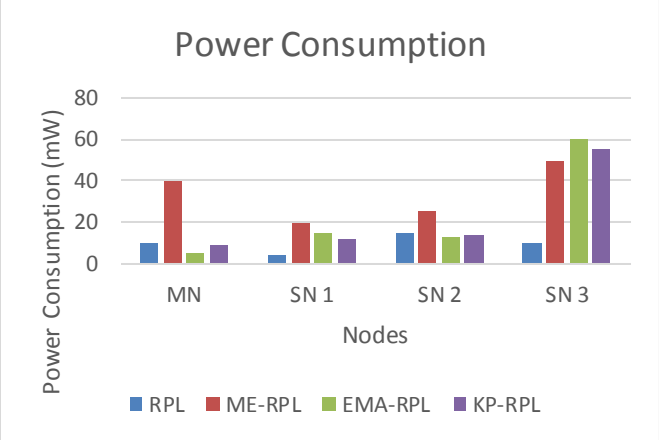


Figure 2. Energy Consumption

A very unique approach is used in [14] which uses corona to deal with mobility issue. In this, the author have not applied any constraint that a mobile node need to be a leaf node. A mobile node can also be a parent node. Some fields are added for Co-RPL in default RPL control messages to integrate mobility with minor modification. There are many DAG roots which are considered to be static and DAG building starts with them. Each DAG root forms a corona. The corona ID of root node is selected as zero. Upon receiving DIO message, the node selects the best parent using lowest corona ID, then determine its own corona ID by adding up one in the lowest corona ID of the received neighbor's message. After calculating the corona ID, it updates the DIO, and broadcast it to neighbor nodes. When due to mobility a node is unable to reach its destination then data packet is send to any node with upper corona ID and the parent node ask its children node to stop sending data. In case of receiving a DIS message before expiry of trickle timer, DIO message is sent immediately before the expiration of trickle timer. The advantage of this mechanism is the packet loss ratio reduces to a considerable number with low energy consumption. Since it allows the parent node to be mobile, route instability increases which not only affect its own data but other child node's data also.

In order to provide mobility support to networks in which data traffic varies highly, the author in Backpressure RPL (BRPL) [15] proposed the merging of RPL and the concept of backpressure routing. BRPL allow multiple logical DAG to be created to handle mobility and traffic. To achieve the desired result for each DAG, a buffered packet queue is maintained by each BPRPL node. Since there are many DAG there are many routes from one source to destination hence when a route is disconnected due to mobility or when there is congestion in a particular route another route can be followed to improve QoS. In order to handle mobility a parameter called theta θ is used whose value ranges between zero to one. The value of this parameter is adjusted dynamically by using QuickTheta algorithm that calculates the θ value based on parameters like congestion level and the mobility of nodes. The advantage is that it reduces the packet loss but it increases the end-to-end delay.

Another approach similar to the approaches using only fixed node as parent node is used in [16]. The author unlike traditional RPL which selects the parent node on the basis of least rank and Euclidean distance, also considers other parameter for the selection of parent node. A preferred parent node set is first constructed using node rank and Euclidean distance (priority given to node rank) , and then among the preferred parent list the best parent is selected using various other parameters obtained from DIO message. In order to look for mobility a D-trickle algorithm is used in which the trickle timer is dynamically calculated using neighbor nodes. It is a simple and efficient algorithm to deal with mobility. It supports mobility with no change in RPL control message and increases packet delivery ratio. The disadvantage is it may not give satisfactory end to end delay and it is reactive protocol.

In order to manage downward traffic under mobility the author in [17] assumes that the new parent selected would be any neighbor node only. Hence in the absence of ACK the parent node broadcast the data destinationed to mobile node to all the neighboring nodes. The new parent selected will transfer the data without delay. This algorithm definitely improves downward traffic management but it has a drawback that broadcasting effects the energy consumption of network and also causes unnecessary flooding of messages. The assumption that the new parent node would be a neighbor node is also not always true.

A route stable strategy is presented in [18] called KP-RPL. It is also a proactive protocol and like any other of this type it

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

also differentiate between mobile nodes and static nodes. The mobile node is never allowed to be a parent node. A list of blacklist nodes is created by all the mobile nodes, which includes static nodes having low RSSI value and can possibly produce unreliable routes. The route produced in this algorithm are quite stable. However the assumption of low mobility make this protocol restricted to some areas only. The energy consumption of mobile node is also high as most of the calculations are performed by mobile node itself and even sometimes positions are not accurately calculated.

An integration of RPL with smart hop is presented in mRPL [19]. The author have proposed a solution to enhance mobility in RPL by actively reacting to any disconnection in the network. The algorithm is broadly divide into two phases as data transmission phase and Discovery phase. In data transmission phase a mobile node based upon the value of average RSSI and various timer value decides whether to send data or stop sending data and enter the next phase. In discovery phase, new parent node is detected using ARSSI value. This value is calculated by every neighbor node from the DIS burst broadcasted by mobile node and a unicast DIO message is sent by neighbor node to mobile node. On the basis of ARSSI value received from all the neighboring node, the node with best ARSSI value greater than a threshold is selected as parent node or access point. This protocol can neither be categorized as proactive or reactive protocol as if it comes know about weakening of a link, it immediately stop sending data to avoid any loss and can be called a proactive protocol but if this disconnection is not detected at an earlier stage then it act as reactive algorithm.

In order to improve RPL for mobile environment, the authors in [20] proposes D-RPL, a routing protocol for IoT applications which are dynamic. This algorithm dynamically adjust the trickle time period on the basis RSSI value received from the messages. Hence it deal with mobility in reactive mode. In order to maximize the responsiveness and smooth transitions and minimize handover delay it includes some routing metrics in calculation of trickle timer. This same approach with some modification is adopted in [21]. (GTM-RPL) like any other proactive approach uses received message RSSI value to detect mobility. A game is developed in which nodes act as a player and compete to send data.

An additional layer which is used for routing is introduced in [22]. All the routing messages passes through this layer while other messages containing data do not go through this layer to avoid any overhead. This algorithm uses fuzzy estimator to find the best parent. Information about the neighbors is fed into fuzzy estimator and on the basis of output route is established. The link disconnection is detected by the new layer introduced using the count of retransmissions. The mobility is handled either by reconstructing the default route or by sending a message to non-root destination about disconnection. The described algorithm is reactive protocol which allow a mobile node to be parent node and sometimes handover delay is increased considerable compared to RPL protocol.

To support mobility in healthcare author in [23] proposed a simple adaption of RPL named mod-RPL(modified RPL). This algorithm just avoid mobile node to be a part of routing path. In order to achieve so mobile nodes are not allowed to send a DIO message and hence it can be only leaf node in DODAG tree. It is a very simple protocol that do not take into account some of the important and complex issues like handover delay and route stability. As healthcare applications are time critical, they need a proactive protocol that can find a new connection before getting disconnected. This is a simple approach which regulates the transmission of control messages.

In [24] ,to handle mobility the author has designed a timer algorithm. The preferred parent is selected on the basis of lowest rank which is calculated using ETX (Expected number of transmission) and RSSI value. A value called time to leave(TL) is calculated which gives the amount of time it will require for the mobile node to be out of range of the parent node. Initially DODAG tree is constructed by using default RPL method and then routes are refined by using RSSI value and avoid any mobile node to be a parent node. According to the timer set, the frequency of DIO message is changed.

The proposed algorithm in [25] reduces the power consumption of mobile node by dynamically changing the DIO timer. Compared to RPL it also reduces the handover delay by reducing the DIO timer when a high mobility is detected using a quantity called mobility level(ML). Unnecessary flooding of control message is also reduced when a low change in network topology is observed by increasing the DIO timer. Though it successfully reduces power consumption effectively but is a reactive protocol. Data loss might also occur due to this.

A demo model [26] is proposed which uses four mechanism to support mobility in RPL. First is the differentiation between mobile node and fixed node. The second is to avoid selecting disconnected routing path by adaptive timing & probing. The third uses two Metrix as RSSI and hop distance for parent selection and the fourth is proactive neighbor discovery so that a new route is found when needed. This algorithm is a combination of some of the popular mobility enhanced RPL. It would have been much better if the proposed method have also included route stability and handover delay.

An algorithm similar to the algorithm in [11] is described in [27]. The entire process is almost the same with some differences. In the later one for parent selection three parameters are used as ETX, RSSI, residual energy. The new parent selection steps in this algorithm is performed by mobile node itself after the mobility is detected using DIS message and hence consumes more power of mobile node. Route stability issue exists in this algorithm also.

A content Centric based routing is proposed in [28]. The algorithm is roughly divide into two parts. The first part deals with mobility enhancement which is carried out by assuming that the new parent node selected would be neighbor of old parent node. Hence when a connection is lost the data for mobile node is broadcasted to all the neighbor node. The second part introduces content centric approach for power

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

consumption and traffic management. A mobile nodes is distinguished from other static nodes using a flag. The proposed algorithm is reactive in nature and the assumption that new parent node is always a neighbor of old parent node give unstable routing path.

IV COMPARISON AND CLASSIFICATION

In this section we analyze various protocols. Table 1 gives the limitations of each discussed algorithm and whether a mobile node is allowed to be a parent node or not.

Table I. Comparison of discussed protocols

Protocol	**Limitations**	**Mobile node allowed to be non leaf node**
ME-RPL [9]	Increased handover delay	No
EMA-RPL[11]	Unstable routes	No
EKF-MRPL[13]	Consume more energy of mobile node	No
Co-RPL[14]	Allow mobile node to be parent and hence increases route instability	Yes
BRPL[15]	Increased end to end delay	Yes
MAEEPS-RPL[16]	Reactive and increased end to end delay	No
KP-RPL[18]	Mobile node resources are wasted.	No
D-RPL[20]	Reactive and unstable routes	Yes
MoMoRo[22]	Increased handover delay	Yes
Mod-RPL[23]	Handover delay is not considered.	No
Adaptive Timer RPL[24]	Increased flow of control messages.	Initially yes, later No
Improved Power Consumption[25]	Date loss.	Yes
Demo model[26]	No new improvement.	No
New method to improve RPL[27]	Route instability and consumes more resources of mobile node.	No
Content Centric Routing In RPL[28]	Assumes new parent node will be a neighbor node.	-

Some of the proposed protocol like those described in [17] and [28] integrate one part of the mobility which manages downward traffic in DODAG. Other described protocol do not look into this aspect. Algorithm in [14] and [22] also somewhat deal with downward traffic but to some extent only. Algorithms in [13], [15] and [18] provide more route stability compared to other algorithms as they select the new parent node based on some conditions and metrics.

Most of the work studied in this survey can be categorized based on the way mobility is handled. Broadly most of the methods can be classified as either proactive or reactive protocol. Some methods which deal with downward traffic are not categorized among these two categories because of the entirely different approach and goal. Further most of the reactive protocol rely on trickle timer to handle mobility while some can be categorized into miscellaneous or other types which do not change the trickle timer duration. Figure 3 shows this categorization of the studied protocols. mRPL is not categorized in these categories as its behavior changes according to received information.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

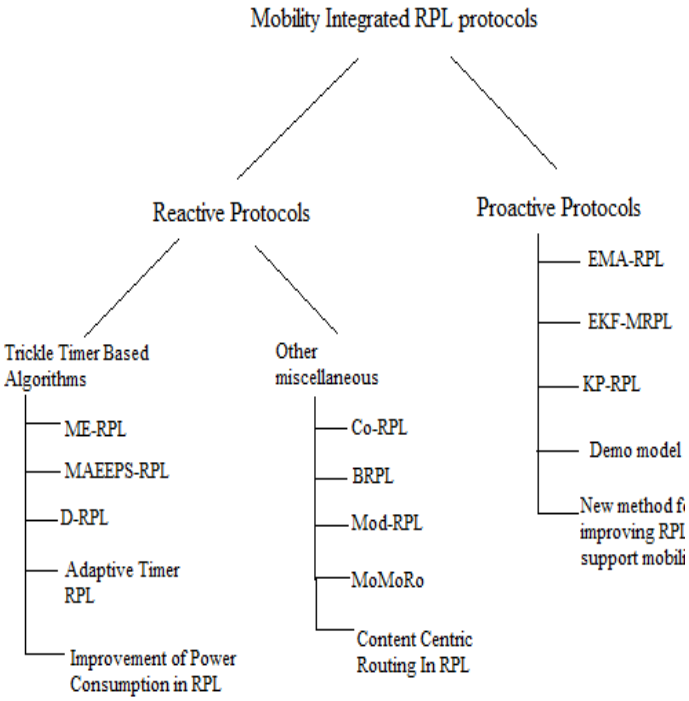


Figure 3. Categorization of discussed algorithms

V CONCLUSION

Mobility is a revolutionary change in IoT which have brought a great enhancement with certain issues. As discussed various authors have successfully resolved those issues. More work can be done to increase route stability with low handover delay and less data loss. WSN are future for healthcare and military services and enhancement can lead to revolutionary changes and ease of life.

REFERENCES

1. P. López, D. Fernández, A. J. Jara and A. F. Skarmeta, "Survey of Internet of Things Technologies for Clinical Environments," International Conference on Advanced Information Networking and Applications Workshops, Barcelona, July 2013.

2. M. K. Singh, S. I. Amin, S. A. Imam, V. K. Sachan and A. Choudhary, "A Survey of Wireless Sensor Network and its types," 2018 International Conference on Advances in Computing, Communication Control and Networking, pp. 326-330 , July 2019.

3. H. Kim, J. Ko, D. E. Culler and J. Paek, "Challenging the IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL): A Survey," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2502-2525, Sept 2017.

4. L.Bartolozzi, F.Chiti, R.Fantacci, T.Pecorella and F.Sgri lli, "Supporting monitoring applications with mobile Wireless Sensor Networks: The eN Route forwarding approach", IEEE International Conference on Communications pp. 5403-5407, Ottawa, Canada, June 2012.

5. V. Safdar, F. Bashir, Z. Hamid, H. Afzal, and J. Y. Pyun, "A hybrid routing protocol for wireless sensor networks with mobile sinks," in Proc. 7th International Symposium Wireless Pervasive Computer, Jul. 2012

6. JP Vasseur, Navneet Agarwal, Jonathan Hui, Zach Shelby, Paul Bertrand and Cedric Chauvenet, "RPL: The IP routing protocol designed for low power and lossy networks", IPSO alliance, April 2011.

7. P. Thubert, "Objective function zero for the routing protocol for lowpower and lossy networks (RPL)," Internet Eng. Task Force, RFC 6552, vol. 20, Mar. 2012.

8. P.Levis, T.Clausen, J.Hui, O.Gnawali, J.Ko, "The Trickle Algorithm", IETF Request for Comments 6206, March 2011.

9. I. E. Korbi, M. Ben Brahim, C. Adjih and L. A. Saidane, "Mobility Enhanced RPL for Wireless Sensor Networks," International Conference on The Network of the Future, pp. 1-8, Feb 2013.

10. M. Bouaziz, A. Rachedi, "A survey on mobility management protocols in Wireless Sensor Networks based on 6LoWPAN technology", Elsevier Journal of Computer Communications, Volume: 74, pages: 3–15, Oct 2016.

11. M. Bouaziz, A. Rachedi, A. Belghith,M. Berbineau and S. Al-Ahmadi, " EMA-RPL: Energy and mobility aware routing for the Internet of Mobile Things", Future Genereration Computer System, vol.97, pp. 247–258, Feb. 2019.N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," in IEEE Signal Processing Magazine, vol. 22, no. 4, pp. 54-69, July 2005.

12. Md.S. Hossen, A.F.M.S. Kabir, R.H.Khan and A. Azfar, "Interconnection between 802.15.4 Devices and IPv6: Implications and Existing Approaches", IJCSI, Vol.7, Iss 1, N.1, Jan 2010.

13. M. Bouaziz, A. Rachedi and A. Belghith, "EKF-MRPL: Advanced mobility support routing protocol for internet of mobile things: Movement prediction approach.", Future Genereration Computer Syststem, Dec. 2017.

14. O. Gaddour, A. Koubäa, R. Rangarajan, O. Cheikhrouhou, E. Tovar and M. Abid, "Co-RPL: RPL routing for mobile low power wireless sensor networks using Corona mechanism," IEEE International Symposium on Industrial Embedded Systems, pp. 200-209 Aug 2014.

15. Y. Tahir, S. Yang and J. McCann, "BRPL: Backpressure RPL for High-Throughput and Mobile IoTs," IEEE Transactions on Mobile Computing, vol. 17, no. 1, pp. 29-43, 1 Jan. 2018.

16. S. Murali and A. Jamalipour, "Mobility-Aware Energy-Efficient Parent Selection Algorithm for Low Power and Lossy Networks," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2593-2601, April 2019.

17. S. Min, S. Chung and Y. Ha, "An Improved Mobility Support Mechanism for Downward Traffic in RPL," International Conference on Ubiquitous and Future Networks, pp. 311-313, Aug 2018.

18. M. Barcelo, A. Correa, J. L. Vicario, A. Morell and X. Vilajosana, "Addressing Mobility in RPL With Position Assisted Metrics," IEEE Sensors Journal, vol. 16, no. 7, pp. 2151-2161, April1 2016

19. H. Fotouhi, D. Moreira, M. Alves, and P.M. Yomsi, "mrpl: Boosting mobility in the Internet of Things", Ad Hoc Networks, vol.26, pp.17-35, 2015.

20. C. Kiraly, T. Istomin, O. Iova and G. P. Picco, "D-RPL: Overcoming memory limitations in RPL point-to-multipoint

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

routing," IEEE Conference on Local Computer Networks, pp. 157-160, Jan 2016
21. H. Kharrufa, H. Al-Kashoash and A. H. Kemp, "A Game Theoretic Optimization of RPL for Mobile Internet of Things Applications," IEEE Sensors Journal, vol. 18, no. 6, pp. 2520-2530, March 2018
22. J. Ko and M. Chang, "MoMoRo: Providing Mobility Support for Low-Power Wireless Applications," in IEEE Systems Journal, vol. 9, no. 2, pp. 585-594, June 2015
23. F. Gara, L. Ben Saad, R. Ben Ayed and B. Tourancheau, "RPL protocol adapted for healthcare and medical applications," International Wireless Communications and Mobile Computing Conference, pp. 690-695, Aug 2015
24. F. Gara, L. Ben Saad, E. Ben Hamida, B. Tourancheau and R. Ben Ayed, "An adaptive timer for RPL to handle mobility in wireless sensor networks,"International Wireless Communications and Mobile Computing Conference, pp. 678-683, Sept 2016
25. Y. Shin and S. Seol, "Improvement of Power Consumption in RPL-based Networks for Mobility Environment", International Conference on Electronics, Information, and Communication, pp. 1-3, April 2020.
26. H. Kim, J. Youn, H. Kim, S. Yoon and S. Bahk, "Demo: Mobility Enhanced RPL for General Mobility Scenarios," IEEE Wireless Communications and Networking Conference Workshops pp. 1-2, June 2020
27. S. Hoghooghi and R. Javidan, "Proposing a new method for improving RPL to support mobility in the Internet of things", IET Networks, vol. 9, no. 2, pp. 48-55, March 2020.
28. A. Kumar and N. Hariharan, "Enhanced Mobility Based Content Centric Routing In RPL for Low Power Lossy Networks in Internet of Vehicles," International Conference on Intelligent Autonomous Systems (ICoIAS), pp. 88-92, April 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:47:56 UTC from IEEE Xplore. Restrictions apply.

AIP Conference Proceedings

A theoretical comparison between $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ based solar cells

Cite as: AIP Conference Proceedings **2265**, 030609 (2020); https://doi.org/10.1063/5.0017201
Published Online: 05 November 2020

Gagandeep, Mukhtiyar Singh, Ramesh Kumar, Vinamrita Singh, and Sunita Srivastava

View Online Export Citation

Meet the Next Generation of Quantum Analyzers
And Join the Launch Event on November 17th

Register now
Zurich Instruments

AIP Publishing

© 2020 Author(s).

A Theoretical Comparison between CH3NH3PbI3 and CH3NH3SnI3 based Solar Cells

Gagandeep[1, a)], Mukhtiyar Singh[2], Ramesh Kumar[1, b)],Vinamrita Singh[3] and Sunita Srivastava[1,4)]

[1]*Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar, India*
[2]*Department of Applied Physics, Delhi Technological University, Delhi, India*
[3]*Department of Applied Sciences and Humanities, Ambedkar Institute of Advanced Communication Technologies & Research, Delhi, India*
[4]*Department of Physics, Punjab University Chandigarh, India*

Corresponding author: a) gagankundu22@gmail.com
b) rameshkumar.bibiyan@gmail.com

Abstract. The perovskite materials based photovoltaic cells play a crucial role in the conversion of sunlight into electric energy. The commonly used perovskite materials are $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$. In the present manuscript, a theoretical comparison is done by simulating the structures ITO/TiO_2/$CH_3NH_3PbI_3$/p-Gr/Ag and ITO/TiO_2/$CH_3NH_3SnI_3$/p-Gr/Ag structures in AFORS-HET software under the illumination condition of AM 1.5G. Here, TiO_2 and p-Graphene (p-Gr) play the role of electron transporting material and hole transporting material respectively with ITO and Ag as front and back contacts. The thickness of layers is directly related to the cost of the cells. Therefore, we optimized the $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ layers thickness to observe the performance of the device. The power conversion efficiency of 10.62% is obtained for $CH_3NH_3SnI_3$based solar cell at optimized thickness of 800 nm as compared to 9.74% efficiency for $CH_3NH_3PbI_3$based solar cell at an optimized thickness of 700 nm. The textured nature of front surface also enhances the efficiency 10.56% and 12.38% for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ respectively.

INTRODUCTION

The largely employed perovskite material is methyl ammonium lead halide perovskite i.e. $CH_3NH_3PbI_3$ for the perovskite solar cells (PSC) [1, 2]. The PSC possess a heterojunction architecture with three layers [3] i.e ETM (electron transporting material), perovskite absorbing layer and HTM (hole transporting material) with front and back contacts as shown in fig. 1a. The commonly employed ETM and HTM are TiO_2 and organic material spiro-OMeTAD respectively [3]. The synthesis of organic material spiro-OMeTAD is expensive process and enhances the cost of device. Therefore, the researchers are developing different types of HTM in PSCs [4-6].

Graphene can be employed as a novel material in PSCs due to its appropriate electrical and optical characteristics [7]. The graphene can be inserted as ETM, HTM or transfer electrode in PSCs because its work function as well as energy band gap can be tuned under the influence of doping [7-9].

The toxicity of lead in PSC is a major obstacle in its development for commercial application. Therefore, tin based PSC (TPSC) [10] is being explored as a substitute of lead based PSC in future. Keeping this in mind, the

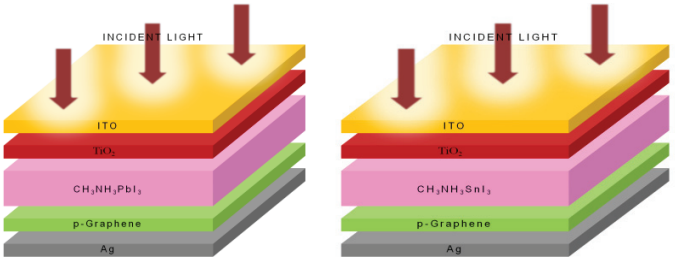
DAE Solid State Physics Symposium 2019
AIP Conf. Proc. 2265, 030609-1–030609-4; https://doi.org/10.1063/5.0017201
Published by AIP Publishing. 978-0-7354-2025-0/$30.00

present work compares the performance of lead-based and lead-free perovskite solar cells in order to develop environmentally friendly technology.

SIMULATION DETAILS

The structures ITO/TiO_2/$CH_3NH_3PbI_3$/p-Gr/Ag and ITO/TiO_2/$CH_3NH_3SnI_3$/p-Gr/Ag are simulated in Automat for Simulation of Heterostructures (AFORS-HET) software [11-13] as shown in fig. 1a and 1b respectively. Both the structures are n-i-p type i.e. light is falling from n-side. The electron-hole pairs are generated in the active layers. The generated electrons are collected by the ETM layer due to the built in voltage developed at the interface between the TiO_2 and perovskite layers. Similarly, generated holes are collected by the HTM layer i.e. p-graphene material.

To check the performance of both the devices, we optimized the active layer thickness. The $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$layer thickness are varied in the range of 50-800 nm and 50-700 nm respectively. The variations in open circuit voltage (Voc), short circuit current density (J_{sc}), fill factor (FF) and power conversion efficiency (PCE) of the cell with absorbing layer thickness are observed. The impact of textured front surface is also studied on both the configurations. The detailed device parameters used in the simulation are taken from different theoretical and experimental studies [3, 14-16] and are mentioned in table 1.



(a) (b)

FIGURE.1 n-i-p structure of perovskite solar cell using (a) $CH_3NH_3PbI_3$ and (b) $CH_3NH_3SnI_3$ absorber layer

TABLE 1. Material parameter of each layer

Parameters	TiO_2	$CH_3NH_3PbI_3$	$CH_3NH_3SnI_3$	Graphene
Thickness (nm)	90	50-800	50-700	0.334
Dielectric constant	100	25	6.5	7.1
Electron affinity (eV)	4	3.9	4.17	4.7
Energy Band gap (eV)	3.2	1.5	1.3	0.5
Optical energy band gap (eV)	3.2	1.5	1.3	0.5
Valence band density (cm^{-3})	1×10^{21}	2.5×10^{20}	1×10^{18}	3×10^{19}
Conduction band density (cm^{-3})	2×10^{20}	2.5×10^{20}	1×10^{19}	3×10^{19}
Hole mobility ($cm^2\ V^{-1}\ s^{-1}$)	0.006	50	1.6	1×10^{5}
Electron mobility ($cm^2\ V^{-1}\ s^{-1}$)	0.006	50	1.6	10
Thermal velocity of electron(cm/s)	1×10^{6}	2.4×10^{18}	1×10^{6}	1×10^{7}
Thermal velocity of hole(cm/s)	4×10^{6}	2.4×10^{18}	1×10^{6}	1×10^{7}
Doping concentration(cm^{-3})	5×10^{19}	1×10^{17}	3.2×10^{15}	1×10^{18}
Layer density(g cm^{-3})	2.328	2.328	2.328	2.328

RESULTS AND DISCUSSIONS

With the help of layer parameters listed in table 1, we studied the effect of absorbing layer thickness of both the perovskite on the performance parameters i.e, J_{sc}, V_{oc}, FF and PCE of the solar cell. Fig.2 gives the effect of thickness of $CH_3NH_3PbI_3$ (Pb) and $CH_3NH_3SnI_3$ (Sn) perovskite layers on the different parameters. It is observed that J_{sc} (fig. 2(a)) shows increment in its value with increase in thickness in both the cases ($CH_3NH_3PbI_3$ and

$CH_3NH_3SnI_3$) to a saturation value of 16.99 mA/cm^2 and 19.90 mA/cm^2, respectively. This enhancement results due to comparatively thicker active layer which absorbs more photons. Greater photon absorption creates large electron-hole pair generation in active layer. The better J_{sc} behavior is demonstrated by $CH_3NH_3SnI_3$ based structure.

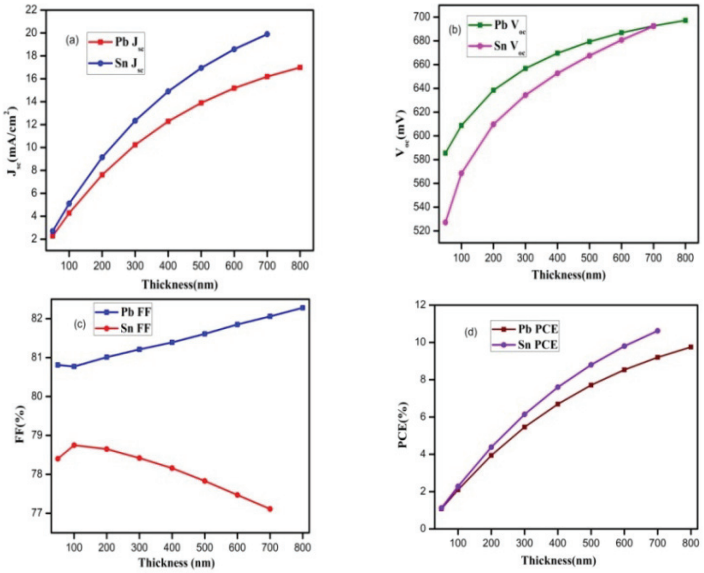


FIGURE 2. Effect of layer thickness of $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ on (a) J_{sc}, (b) V_{oc}, (c) FF and (d) PCE of the perovskite solar cell

The value of V_{oc} varies from 585.5 mV to 697.2 mV and 527.1 mV to 692.2 mV for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ perovskite respectively and the results are given in fig. 2 (b). However, the FF value slightly reduces from 78.40% to 77.11% in $CH_3NH_3SnI_3$ layer within the thickness range of 50-700 nm. This may probably due to internal recombination in perovskite absorber layer. But in $CH_3NH_3PbI_3$, the value of FF increases from 80.81% to 82.28% with thickness ranging from 50-800 nm and the results are given in fig. 2(c). Fig. 2(d) gives the variation of PCE with thickness for both the perovskite solar cells and a considerable increase in the value of efficiency from 1.07% to 9.74% and 1.12% to 10.62% is noted for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ respectively. The performance for $CH_3NH_3SnI_3$ based device is better as compared to $CH_3NH_3PbI_3$based device.

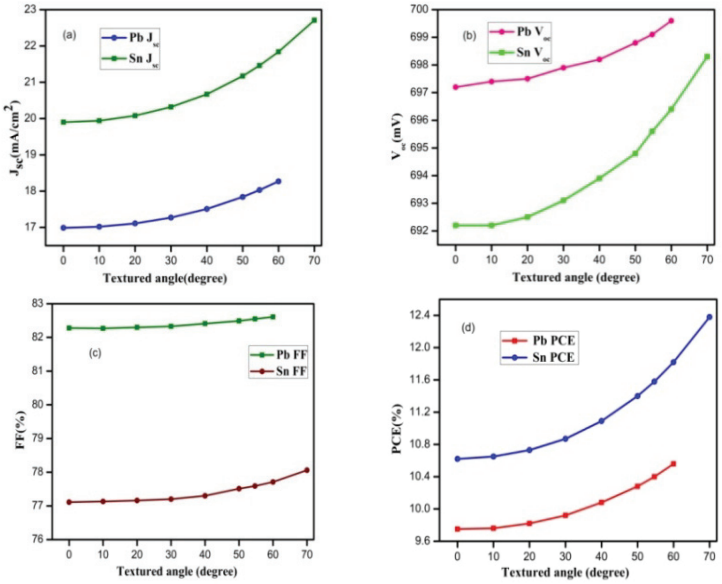


FIGURE 3. Effect of angle of incidence of light on $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ on (a) J_{sc}, (b) V_{oc}, (c) FF and (d) PCE of the perovskite solar cell

The performance of solar cell is also influenced by the nature of the front contact surface because it may be plane or textured. The textured front surface absorbs more photons as compared to the plane front surface. This is due to the fact that when incident light falls on the plane surface of the cell, it may bounce back but if the surface is textured the light undergoes multiple reflections. Hence, in effect, large area is used for the absorption of incident light improving the cell performance. We varied the angle of incidence from 0° to 60° and 0° to 70° for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ respectively and the results are shown in fig. 3. It is observed that all the cell parameters improve with incidence angle giving J_{sc} values 16.99 mA/cm^2 to 18.27 mA/cm^2 and 19.90 mA/cm^2 to 22.71 mA/cm^2, V_{oc} = 697.2 mV to 699.6 mV and 692.2 mV to 698.3 mV, FF = 82.28% to 82.61% and 77.11% to 78.06% and PCE varying from 9.74% to 10.56% and 10.62% to 12.38% for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ perovskite solar cell.

CONCLUSIONS

The present manuscript is devoted to the comparison between the $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$based solar cells. The behavior of both the cells is discussed under the influence of perovskite layer thickness. We also replaced the complex material spiro-OMeTAD by graphene as hole transporting material. The PCE for both $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ solar cells are 9.74% and 10.62% at optimized active layer thickness of 800 nm and 700 nm respectively. The textured nature of front surface also improved the efficiency to 10.56% and 12.38% for $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$solar cells respectively. Therefore, based on the values of performance parameters for both $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ solar cells, the tin based solar cells has found to exhibit better performance for the above suggested structure.

ACKNOWLEDGMENT

One of the authors (**Ramesh Kumar**) is grateful to the **University Grants Commission, INDIA** for providing the financial support through START-UP GRANT under FRPS (No. F.30-364/2017 (BSR), Dated 15/12/2017). The authors gratefully acknowledge the AFORS-HET software developer for providing free access of the software.

REFERENCES

1. A. Kojima, K. Teshima, Y. Shirai and T. Miyasaka J. Am. Chem. Soc. **131**, 6050 (2009).
2. N. J. Jeon et. al., Nat. Energy **3**, 682 (2018).
3. H. J. Du, W. C. Wang and Y. F. Gu, Chin. Phys. B **26**, 028803 (2017).
4. K. Mahmood, B. S. Swain and H. S. Jung, Nanoscale.**6**, 9127 (2014).
5. S. Lu et. al., Chem. Commun.**50**, 6931(2014).
6. P. Qin et. al., Nat. Commun.**5**, 3834 (2014)
7. E. L. Lim et.al., Nano-Micro Lett. **10: 27** (2018).
8. T. Mahmoudi, Y. Wang and Y. B. Hahn, Nano Energy.**47**, 51 (2018).
9. M. F. Bhopal et.al., J. Mater. Chem. C. **5**, 10701 (2017).
10. MH. Kumar et.al., Adv. Mater. **26**, 7122 (2014).
11. R.Varache et. al., Sol. Energ. Mat. Sol. C.**141**, 14 (2015).
12. R. Stangl, J. Haschke and C. Leendertz,Solar Energy, Intech (R.D. Rugescu (Ed.), Croatia, 2009), pp. 432.
13. R.Stangl, M. Kriegel and M. Schmidt, Proceeding of the WCPEC-4, Fourth World Conference on Photovoltaic Energy Conversion (Hawaii, USA), PP. 1350 (2006).
14. S.M. Iftiquar, J. S. Kim and J. Yi, Optik**148**, 54(2017).
15. K. Patel and P.K. Tyagi. Carbon**116**, 744 (2017).
16. Gagandeep, M. Singh and R. Kumar, AIP Conference Proceedings, **2115**, 030548 (2019).

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Accelerating Business Growth with Big Data and Artificial Intelligence

Awishkar Ghimire[1,#], Surendrabikram Thapa[1], Avinash Kumar Jha[2], Surabhi Adhikari[1] and Ankit Kumar[2]

[1]Department of Computer Science and Engineering, Delhi Technological University, India
[2]Department of Civil Engineering, Delhi Technological University, India
[#]awishkar.ghimire@gmail.com

ABSTRACT

Artificial Intelligence (AI) is considered to be the fourth industrial revolution. Artificial Intelligence with the help of big data has transformed all industries around the world. Artificial intelligence refers to the simulation of human or animal intelligence in computational systems so that they are programmed to think like intelligent beings and mimic the actions of intelligent entities. Computational systems which have programmed intelligence can solve different real-world problems far more accurately and efficiently than computational systems that are deterministic and hardcoded. Since many problems in business and business analytics cannot be solved by deterministic systems, AI plays a major role in tackling problems in the business world. Machine learning and deep learning which are subsets of the field of AI is widely used to solve and optimize many problems in business such as marketing, credit card fraud detection, algorithmic trading, customer service, portfolio management, product recommendation according to the needs of customers, insurance underwriting. AI and big data have revolutionized the business world and this paper discusses some AI and big data technologies that are currently being used to accelerate business growth.

Keywords— Artificial Intelligence, Big Data, Business Analytics, Decision Making

1. INTRODUCTION

A business can be defined as an enterprising entity or an organization consisting of people and assets, that is involved in professional, commercial or industrial activities to make a monetary profit. A business can be considered as the backbone of the modern economy. The size of a business can range from a small industry that offers its services in a small town to a very large group of industries that is located in many different countries. Businesses can be owned by a single person as well as many thousands of people [1]. There are many types and forms of businesses and all of them have been affected by modern technology especially the larger corporations. In fact, many large companies like Google, Facebook, Amazon are at a war of technology.

Technology in today's modern world is rapidly improving and is making a huge impact in all sectors of the modern world [2]. It has impacted on medical diagnosis, business and many vital sectors [32]. Many businesses in today's world have started using modern cutting edge technology to accelerate their growth and to skyrocket their profits. Artificial Intelligence, Data Science, Big Data, Internet Of Things (IoT) have completely transformed business environments and the way people do business. In the present day, there is not a single field of work that has not explored the use cases of AI [27]. The uses of AI and different computational technologies being used in the manufacturing industries as well as technology industries can be seen. It has been well documented that machines can outperform or at least match human beings in a wide range of activities including emotion-sensing, tacit judgement and automation activities [31]. It has been estimated that computational technologies could take over as much as 47% of the current jobs in the world in as little as 10 years [3]. Technology has had a huge impact in the economic world and the rapid developments in technology in the upcoming years will have completely change the landscape of the current economy. So, it is paramount to have an understanding of how the current cutting edge technologies are impacting the business world. AI is not only being used to automate tedious manual work but it has also been successfully used to do jobs that require higher-order creative thinking such as the works of journalists, attorneys, lab technicians, paralegals etc. [3]. Due to AI and other computational technologies, many jobs are replaced by computational technologies, but this is only for low-skilled jobs such as clerical work. The demand for high skilled jobs and jobs that require the creation of such computational systems have increased and is speedily increasing. Soni et al. [4] in their paper analyze 100 AI startups in the world and describe a very interesting finding. The total investments in the 100 companies they investigated were $25.88 million in 2011 which increased to $1866.6 million

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

in 2016. In just 6 years the funding in these companies increased by a whopping 7112.52%. These figures give an idea of how artificial intelligence and big data is stirring the current business scenario.

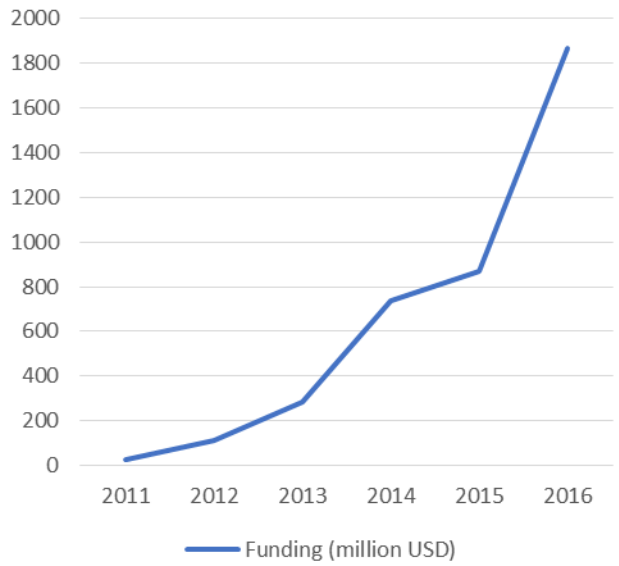


Fig.1. Bar graph of the total investment every year of the 100 AI startups as analysed by Soni et al. [4]

Artificial intelligence and big data are being used in various aspects of businesses to improve profits or to speed up laborious work and thus accelerate business growth. Many works that would have taken thousands of hours for a human to complete can be done in a matter of minutes by an automated system. Many companies like cognizant develop artificially intelligent systems to improve business processes and revenue [5]. In one case the software company Cognizant developed a solution of optimum equipment utilization using data analysis for a global mining company that saved $30 million in the capital due to higher availability of equipment. Thus can be seen that AI and data analysis is being heavily used in the modern industry to cut costs and optimize profits for the business thus accelerating growth.

2. RELATED WORKS

The major areas in finance and business where artificial intelligence and big data solutions are playing a role are loan/insurance underwriting, fraud detection, customer service, sentiment/news analysis, algorithmic trading, portfolio management, marketing, recommendation systems of financial products, advertisement recommendation etc. Some of the AI techniques and their uses in these areas of business are discussed below.

A. Fraud Detection

Fraud is one of the most substantial crises in today's economic world. Businesses all around the world lose billions of dollars due to fraudulent activity. According to various sources, the total fraud losses reached $27.85 billion in 2018 and this figure has been estimated to rise to $40.63 billion in the coming ten year period [6]. This figure is more than the annual GDP of many third world countries. Therefore, computational systems that are capable of detecting and catching fraudulent activities is quintessential, because doing so would rapidly accelerate business and economic growth. Researchers have devised many systems that incorporate artificial intelligence and machine learning to create such systems. All fraud detection techniques can generally be classified into two categories, anomaly detection and misuse detection [7]. Anomaly detection learns the transaction behaviour of a certain customer and any new transaction made by the customer is classified as normal or anomalous according to past transactions made by that customer. Misuse detection creates the model using a labelled data set of all the customers, and fraudulent activity is determined according to the general fraudulent patterns.

Awoyemi et al. [8] compare traditional machine learning models for credit card fraud detection. A dataset consisting of 284,807 credit card transactions from European cardholders was used. Since fraud dataset is highly skewed, oversampling, undersampling techniques along with no sampling was also used for better and more accurate comparison. Their paper compared naive Bayes classifier, logistic regression classifier and k-nearest neighbour classifier [9]. All of these are misuse detection techniques. The naive Bayes classifier gave 97.92% accuracy, the logistic regression classifier gave 54.86% accuracy and the k - nearest neighbour classifier gave 97.69% accuracy. Logistic regression gives much lower accuracy and this is apparent because it cannot learn non - linear functions.

Xuan et al. [7] in their paper mention a random forest algorithm for credit card fraud detection. The dataset used by them comes from a Chinese e-commerce company. Random forest is an ensemble learning method in which a set of decision trees is created, and every tree is made with an independent data set. The performance of the random forest depends on the strength of each tree and the correlation between the trees [28]. The higher the strength of each tree and the lower the correlation between the trees, the better the performance of the algorithm. Xuan et al use two different random forest models and compare their performance on the dataset. The two random forests differ in their base classifier i.e individual tree in the model. The first random forest uses a simple decision tree as the base classifier and the second random forest uses CART(classification and regression) trees as the base classifier. The accuracy of the first random forest comes out

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

to be 91.96% and the accuracy of the second model comes out to be 96.77%.

Randhawa et al. [10] compare the performance of about 12 machine learning algorithms on credit card fraud detection. They use a public credit card transaction dataset. They also use ensemble methods which use adaptive boosting and majority voting for classifying an instance. They experiment on 7 ensemble methods which use majority voting. The best accuracy was given by an ensemble of Neural Network and Naive Bayes classifier. This ensemble had an accuracy of 99.941% on the dataset they used. In the real world, companies generally use ensemble methods for fraud detection.

B. Algorithmic Trading

Trading is one of the key concepts in economics and business. Trading can loosely be defined as the buying and selling of an economic entity such as goods, stocks, currency etc. Trading is undertaken by individuals and businesses to earn a profit. The trading process generally consists of 4 components, pre-trade analysis, trading signal generation, trade execution and post-trade analysis [11]. Algorithmic trading can be loosely defined as the automation of any of the combination of these steps or all of these steps. Artificial Intelligence has completely transformed this area of business by automating the trading process and many trading algorithms can create profit without any human intervention. Algorithmic trading using artificial intelligence has had a huge impact on business and growth.

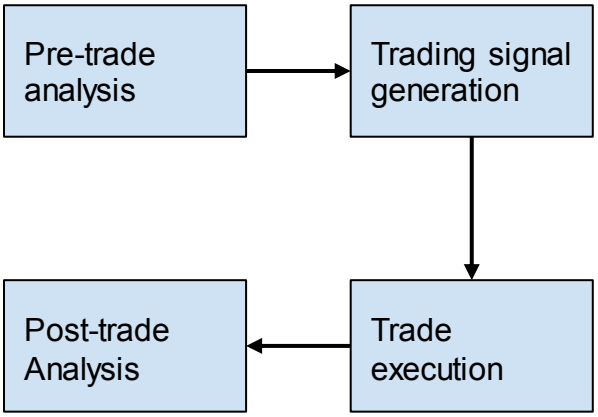


Fig. 2. Steps in a complete algorithmic trading system

Roondiwal et al. [12] propose a learning algorithm that can be used to predict stock prices. The dataset that they use is of New York stock exchange and the dataset consists of the data, the open price, the close price, the volume. The use LSTM (Long Short Term Memory) which is a type of recurrent neural network. LSTMs incorporates a memory cell which corresponds to neurons in traditional artificial neural networks. These memory cells can associate memories in the input and grasp the structure of data and hence make the prediction accurate. The LSTM they used consisted of a sequential input layer after which comes the 2 LSTM layers and a dense layer which uses ReLU activation function and then finally the output layer which uses the linear activation function. Roondiwala et al conducted various experiments tweaking different parameters and the best performing LSTM displayed an RMSE(Root mean square error) of 0.00859 which is unbelievably low and hence proved that Artificial intelligence can be effectively used to predict stock prices.

Colianni et al. [13] use sentiment analysis to predict the prices of bitcoins. They use an open-source API to gather the tweets with the keyword bitcoin. Then the features are extracted from the tweet in the form of words and their counts. This feature then becomes an input to the classification algorithm which tells the bitcoin price will increase or decrease. The classification algorithms mentioned in their paper are naive bayes classifier and support vector machines. The SVM gives better accuracy according to their paper. Their algorithm is limited in the sense that it doesn't tell what the increasing percentage of the decrease percentage will be, however, bitcoin price prediction using sentiment analysis in twitter is an interesting application of artificial intelligence.

C. Customer Service

Artificial Intelligence and data science solutions have also been used to automate customer service [14]. With AI technology businesses no longer need to hire dedicated customer service officials. A 24/7 chatbot can answer the different queries of customers and this practice is seen widely successful in the last few years in different businesses. AI can handle questions of the customers from a huge knowledge database it has learnt from. This substantially decreases the hiring cost of a company, and the same work can be done with much less money, thus empowering the growth of the business.

Cui et al. [15] developed a chatbot called superagent that acts as a customer service provider for e-commerce websites like amazon. It is an add on and it can work on any e-commerce website. The system first crawls the page of the product that the user searches for and gathers the information from the page. The system uses state of the art NLP techniques to interpret the question that the user asks it. To find information regarding the subject matter asked the system uses deep learning models and recurrent neural networks and further uses NLP to answer the questions in English. Super agent system has 3 engines that are effectively used to answer the questions of the customers. The first engine is FACT QA that is used to answer questions related to the product information like "what is the specification of the CPU". The second engine is FAQ search

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

for customer QA pairs. This engine is used to answer questions like "does it come with ink" etc. The third engine is Opinion-Oriented Text QA used for reviews. This engine is used to answer questions like "what do you think of the design of the cup".The system is also capable of answering small talk questions such as hello etc. Systems like these can potentially save millions of dollars for a company.

D. Marketing and Product Recommendation

Artificial Intelligence and big data have had a considerable impact on marketing as well. Data-driven marketing strategies are much more effective than using human-based marketing strategies. These days youtube. Facebook, Instagram and almost all social media uses artificial intelligence and machine learning systems to display individualized and personalized advertisements. A person that is interested in guitar is much more likely to be shown an advertisement related to guitar in such platforms. Marketing is arguably the most heavily affected area in business due to artificial intelligence. The centrepiece in marketing today is customer analytics using big data and the interpretation of this data using various machine learning models that produce amazing and profitable insights [16].

Sundsoy et al. [17] experimented on an MNO(mobile network operator) in Asia comparing big data-driven marketing strategies to traditional human-based marketing strategies. The company often used to send marketing SMS to selected customers regarding some data packs and hope that the customer would buy it. The customers were selected by marketing officials. In their experiment sundsoy et al divides customers into a control group and treatment group. The control group was selected by the marketing officials of the company whereas the treatment group was selected by using some data-driven machine learning models. After the experiment, they found that the conversion rate i.e the customers who were willing to buy the marketed data pack was 13 times higher in the treatment group(group selected using data-driven models) than the control group(group selected by marketing officials). They tested using different machine learning models such as artificial neural networks, support vector machines, and the final model was a bootstrapped decision tree. This experiment clearly shows the power of data-driven analytics and how it can impact and accelerate business growth. Product Recommendation in websites is also another form of marketing that is widely being used in today's world.

Paradarami et al. [18] propose a hybrid recommender system for recommending business products. Their recommendation system consists of a deep artificial neural network that uses reviews along with content-based features and collaborative features to create accurate recommendations. This paper demonstrates that ANN trained for a specific class of business can generalize into other classes of business as well and recommend products for other classes very effectively as well. It also generalizes into different types of users as well and hence the scalability of the recommendation system is drastically improved. This approach only requires model parameters to be saved in the memory unlike memory-based recommendation systems, this significantly decreasing digital memory footprint and greatly increasing scalability. Their approach can also be used to implement real-time recommendation systems which work in an instant due to the speed of their system.

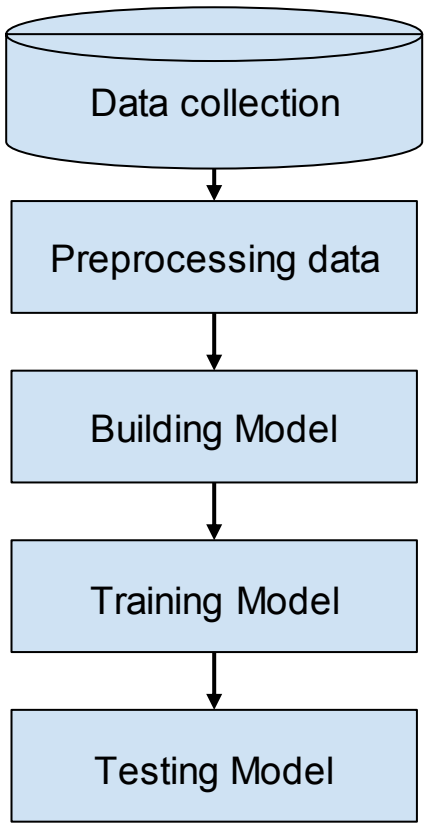


Fig. 3. General scheme of an AI model

E. Cyber Security

Cybercrime is one of the greatest threats to mankind in the modern world. It is the most prominent threat to any modern business in the current day. It has been estimated that the loss due to cybercrimes is projected to reach $6 trillion by 2021, which is staggeringly huge. Technologies and strategies to fight cybercrime are paramount in business. Many businesses have gone bankrupt due to cybercrimes. Effective cybersecurity technology against cybercrimes will invariably help the growth and sustenance of a business [19]. Artificial intelligence and machine learning have been successfully used to fight against cybercrimes. Application of machine learning in cyber crimes is generally categorized into three categories, intrusion detection, malware analysis and phishing detection [20].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

Kim et al. [21] used Long short term memory (LSTMs) which is a type of Recurrent neural network for intrusion detection in cybersecurity systems. The data set they used to conduct their experiments was KDD Cup 1999 dataset. 4,898,431 network traffics were present in the data set and each traffic consisted of 41 features like IP address etc and 22 attacks are differentiated according to their characteristics. In their experiment, they compared many different machine learning algorithms and LSTM-RNN has the highest accuracy of 96.93%. Another application of machine learning in cybersecurity is phishing detection.

Gupta et al. [22] propose an artificial neural network with particle swarm optimization(PSO) to detect phishing URLs. The PSO techniques are used widely because of its robustness and short computational time [29]. Moreover, they also have an easy implementation [30]. In their paper, they compare back propagation neural network and part PSO neural network and they find that in their dataset, PSO

TABLE I. PAST WORK IS DONE IN THE FIELD OF BUSINESS ANALYTICS

Name of Author	Year of Publication	Methods used	Dataset used	Use case
Paradarami et al. [18]	2017	Deep ANN + content-based features + collaborative features	Yelp academic dataset	Product Recommendation
Awoyemi et al. [8]	2017	K-nearest neighour/ naive bayes/ logistic regression	Credit card transaction of european cardholders dataset	Fraud detection
Xuan et al. [7]	2018	Random forest	Credit card transaction dataset of chinese e-commerce company	Fraud detection
Randhawa et al. [10]	2018	Adaboost + majority voting , ANN + NB	Public credit card transaction dataset	Fraud detection
Roondiwala et al. [12]	2017	LSTM	New York stock exchange dataset	Stock prices prediction using time series data
Colianni et al. [13]	2015	NLP + SVM/NB	Twitter data	Bitcoin price prediction using sentiment analysis
Cui et al. [15]	2018	NLP + Deep ANN	E-commerce data	Customer service
Kim et al. [21]	2016	LSTM	KDD cup 1999 dataset	Intrusion detection in cyber security
Gupta et al. [22]	2017	ANN with PSO	UCI repository archive	Phishing detection in cyber security

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

neural network gave an accuracy of 93.93% accuracy when the number of hidden neurons was 67 and the activation function used was tansig. The highest accuracy that the Backpropagation algorithm gave was 96.81% when the number of hidden neurons was 21 and the activation function was logsig. Machine learning is heavily influencing the field of cybersecurity and many companies are using it to fight cybercrimes and save billions of dollars.

3. FURTHER APPLICATIONS AND DISCUSSION

Table I shows the past works that have been done in the last five years or so. The works show that there are ample areas where artificial intelligence can be used to grow business in a lot of ways ranging from fraud detections to product recommendation. The recent use of chatbots in customer service experience was a hugely celebrated success and it shows that there is a great scope of AI in business analytics. Similarly, small stock traders and bitcoin traders also use Business analytics to build predictive models.

In the above section, various papers on artificial intelligence and data science pertaining to the use cases in business and finance have been discussed. However, AI is not just limited to the above-mentioned areas of finance. There are a lot more areas of finance and business in which artificial intelligence and data science have been successfully used to make huge progress in a particular field. AI is increasing the efficiency and productivity across the entire value chain in a business process, right-front the moment the business model is developed to the moment the business product goes to the hands of the customer. AI is revolutionizing the business process.

One interesting application of AI in business technology is automatic retail transactions in supermarkets. Amazon with its Amazon Go retail stores are using this technology. There are recognition sensors of various kinds throughout the store that sense what the customer is picking up int the store and they automatically charge the customer without the customer having to checkout through the traditional process [23]. Another area in which data science and AI has had immense value is credit scoring. AI and ML algorithms have been successfully used to give a credit score to customers given the customer data.

Artificial intelligence has also been successfully used in portfolio management. Portfolio management is the robust decision-making process in which funds are allocated in different types of investment products to reduce the risk of losing too much money. Jiang et al. [24] propose a deterministic deep reinforcement learning method for solving the portfolio management problem when investing in cryptocurrencies. Their approach doesn't rely on any financial theory hence, it is robust and can be extended to any other financial markets as well. Liu et al. [25] propose a one-layer recurrent artificial neural network for solving linear pseudoconvex optimization problems with constraints for dynamic portfolio optimizations to improve profits and reduce the loss of investment opportunities.

A lot of Quantitative Finance is another field of business in which machine learning has played a huge role. Spiegeleer et al. [26] propose various machine learning models for solving traditional quantitative finance problems like hedging, curve fitting, derivative pricing. They have also given a novel idea which used a Bayesian non-parametric technique called the GPR. They use ML models to fit sophisticated greek profiles and summarise different implied volatility surfaces.

According to an article published in the Harvard business review [33], three kinds of artificial intelligence are taking over the business industry. The first kind is process automation. This included reading contractual and legal documents using state of the art NLP techniques and extracting and interpreting it using different intelligent models, transferring relevant data and information from call centre systems and emails to the databases of the company. The second kind is cognitive insights. This kind of system solves problems like what the customer is more likely to buy, analyze different warranty data to correctly identify quality problems or safety in different manufactured problems. provide insurance companies with detailed and more accurate modelling. The third kind mentioned in the article is cognitive engagement. This includes systems like recommendation systems for health treatment that help users get a customized care plan that takes into account each user's health history and different data. In that article, the author's reviewed 152 different use cases of artificial intelligence in business systems, among which 47% fell into the process automation category, 38% fell into the cognitive insights category and 16% fell into the cognitive engagement category.

4. CONCLUSION AND FUTURE WORKS

In this paper, several applications of machine learning, deep learning and data science in the context of business and their growth have been reviewed. AI has utterly transformed the business world, and many people in today's world have been working side by side with artificial intelligence systems to solve complex problems in the world and make a profit out of the solution. AI has been seen used in various areas of the business to accelerate its growth. Some fields which have been discussed in this paper are marketing, customer service, algorithmic trading, fraud detection, cybersecurity, portfolio management, sentiment analysis, credit scoring.

AI has been used successfully to drive the modern economic world and it will be a key player in shaping the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

future. Today's AI is not even as intelligent as human beings yet it has had a massive impact in the world. The AI of tomorrow will undoubtedly change the way the business and how businesses function is seen. Once AI reaches a certain level, might be able to see systems that can hire people into a company. Artificial Intelligent systems might someday take over the work of a CEO and might overlook businesses and how they function. In the future, AI systems might do every job from building a business model to supplying it into the market rendering human intervention obsolete.

REFERENCES

[1] S. Prakash et al., "Characteristic of enterprise collaboration system and its implementation issues in business management," International Journal of Business Intelligence and Data Mining, vol. 16, no. 1, pp. 49-65, 2020.

[2] S. Adhikari, S. Thapa, and B. K. Shah, "Oversampling based Classifiers for Categorization of Radar Returns from the Ionosphere," in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020: IEEE, pp. 975-978.

[3] S. A. Wright and A. E. Schultz, "The rising tide of artificial intelligence and business automation: Developing an ethical framework," Business Horizons, vol. 61, no. 6, pp. 823-832, 2018.

[4] N. Soni, E. Sharma, N. Singh, and A. Kapoor, "Impact of Artificial Intelligence on Business," in Digital Innovations, Transformation, and Society Conference 2018 (Digits 2018). pp, 2018, vol. 10.

[5] A. Kumar and M. Janakiraman, "Cognizant Technology Solutions: Growth and transformation of its data warehousing and business intelligence division," Journal of Information Technology Case and Application Research, vol. 10, no. 3, pp. 56-83, 2008.

[6] S. Rajora et al., "A comparative study of machine learning techniques for credit card fraud detection based on time variance," in 2018 IEEE Symposium Series on Computational Intelligence (SSCI), 2018: IEEE, pp. 1958-1963.

[7] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018: IEEE, pp. 1-6.

[8] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in 2017 International Conference on Computing Networking and Informatics (ICCNI), 2017: IEEE, pp. 1-9.

[9] A. Saxena et al., "A review of clustering techniques and developments," Neurocomputing, vol. 267, pp. 664-681, 2017.

[10] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," IEEE access, vol. 6, pp. 14277-14284, 2018.

[11] G. Nuti, M. Mirghaemi, P. Treleaven, and C. Yingsaeree, "Algorithmic trading," Computer, vol. 44, no. 11, pp. 61-69, 2011.

[12] M. Roondiwala, H. Patel, and S. Varma, "Predicting stock prices using LSTM," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp. 1754-1756, 2017.

[13] S. Colianni, S. Rosales, and M. Signorotti, "Algorithmic trading of cryptocurrency based on Twitter sentiment analysis," CS229 Project, pp. 1-5, 2015.

[14] S. Thapa, S. Adhikari, and S. Mishra, "Review of Text Summarization in Indian Regional Languages," in 2020 International Conference on Computing Informatics & Networks (ICCIN), 2020: Springer.

[15] L. Cui, S. Huang, F. Wei, C. Tan, C. Duan, and M. Zhou, "Superagent: A customer service chatbot for e-commerce websites," in Proceedings of ACL 2017, System Demonstrations, 2017, pp. 97-102.

[16] S. Erevelles, N. Fukawa, and L. Swayne, "Big Data consumer analytics and the transformation of marketing," Journal of business research, vol. 69, no. 2, pp. 897-904, 2016.

[17] P. Sundsøy, J. Bjelland, A. M. Iqbal, and Y.-A. de Montjoye, "Big data-driven marketing: how machine learning outperforms marketers' gut-feeling," in International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction, 2014: Springer, pp. 367-374.

[18] T. K. Paradarami, N. D. Bastian, and J. L. Wightman, "A hybrid recommender system using artificial neural networks," Expert Systems with Applications, vol. 83, pp. 300-313, 2017.

[19] A. Saini, M. S. Gaur, V. Laxmi, and M. Conti, "Colluding browser extension attack on user privacy and its implication for web browsers," Computers & Security, vol. 63, pp. 14-28, 2016.

[20] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in 2018 10th International Conference on Cyber Conflict (CyCon), 2018: IEEE, pp. 371-390.

[21] J. Kim, J. Kim, H. L. T. Thu, and H. Kim, "Long short term memory recurrent neural network classifier for intrusion detection," in 2016 International Conference on Platform Technology and Service (PlatCon), 2016: IEEE, pp. 1-5.

[22] S. Gupta and A. Singhal, "Phishing URL detection by using artificial neural network with PSO," in 2017 2nd International Conference on Telecommunication and Networks (TEL-NET), 2017: IEEE, pp. 1-6.

[23] C. Campbell, S. Sands, C. Ferraro, H.-Y. J. Tsao, and A. Mavrommatis, "From data to action: How marketers can leverage AI," Business Horizons, vol. 63, no. 2, pp. 227-243, 2020.

[24] Z. Jiang and J. Liang, "Cryptocurrency portfolio management with deep reinforcement learning," in 2017 Intelligent Systems Conference (IntelliSys), 2017: IEEE, pp. 905-913.

[25] Q. Liu, Z. Guo, and J. Wang, "A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization," Neural Networks, vol. 26, pp. 99-109, 2012.

[26] J. De Spiegeleer, D. B. Madan, S. Reyners, and W. Schoutens, "Machine learning for quantitative finance: fast derivative pricing, hedging and fitting," Quantitative Finance, vol. 18, no. 10, pp. 1635-1643, 2018.

[27] S. Thapa, P. Singh, D. K. Jain, N. Bharill, A. Gupta, and M. Prasad, "Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease", in 2020 International Joint Conference on Neural Networks (IJCNN), 2020: IEEE.

[28] A. K. Jha, S. Adhikari, S. Thapa, A. Kumar, A. Kumar, and S. Mishra, "Evaluation of Factors Affecting Compressive Strength of Concrete using Machine Learning" in 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), 2020: IEEE.

[29] N. Bansal, R. Gautam, R. Tiwari, S. Thapa, and A. Singh, "Economic Load Dispatch Using Intelligent Particle Swarm Optimization", in 2020 2nd International Conference on Intelligent Computing, Information and Control Systems (ICICCS 2020), 2020: Springer.

[30] N. Bansal, S. Thapa, S. Adhikari, A. K. Jha, A. Gaba, and A. Jha, "Novel Exponential Particle Swarm Optimization Technique for Economic Load Dispatch", 2nd International Conference on Inventive Computation and Information Technologies, 2020: Springer.

[31] S. Thapa, S. Adhikari, U. Naseem, P. Singh, G. Bharathy, and M. Prasad, "Detecting Alzheimer's Disease by Exploiting Linguistic Information from Nepali Transcript", in 2020 27th International Conference on Neural Information Processing, 2020: Springer

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

[32] S. Thapa, S. Adhikari, A. Ghimire, and A. Aditya, "Feature Selection Based Twin-Support Vector Machine for the diagnosis of Parkinson's Disease", in 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), 2020: IEEE

[33] T. H. Davenport and R. Ronanki, "Artificial intelligence for the real world," Harvard business review, vol. 96, no. 1, pp. 108-116, 201

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:49:19 UTC from IEEE Xplore. Restrictions apply.

Journal of Difference Equations and Applications

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gdea20

An adaptive difference scheme for parabolic delay differential equation with discontinuous coefficients and interior layers

Aditya Kaushik & Nitika Sharma

To cite this article: Aditya Kaushik & Nitika Sharma (2020): An adaptive difference scheme for parabolic delay differential equation with discontinuous coefficients and interior layers, Journal of Difference Equations and Applications, DOI: 10.1080/10236198.2020.1843645

To link to this article: https://doi.org/10.1080/10236198.2020.1843645

Published online: 06 Nov 2020.

Submit your article to this journal

Article views: 16

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gdea20

JOURNAL OF DIFFERENCE EQUATIONS AND APPLICATIONS
https://doi.org/10.1080/10236198.2020.1843645

Taylor & Francis
Taylor & Francis Group

Check for updates

An adaptive difference scheme for parabolic delay differential equation with discontinuous coefficients and interior layers

Aditya Kaushik[a,b] and Nitika Sharma[a]

[a]Department of Applied Mathematics, Delhi Technological University, Delhi, India; [b]University Institute of Engineering and Technology, Panjab University, Chandigarh, India

ABSTRACT

A class of singularly perturbed parabolic delay differential equation with discontinuous data is solved numerically. Numerical scheme based on the upwind finite difference method is presented on a specially generated mesh. The proposed numerical method has been analysed for consistency, stability, and convergence. It is found that the method proposed is unconditionally stable, and the convergence obtained is parameter uniform. Numerical illustrations are presented to demonstrate the effectiveness of the technique.

ARTICLE HISTORY
Received 9 April 2020
Accepted 20 October 2020

KEYWORDS
Singular perturbation; delay differential equation; finite difference method; interior/boundary layers

2010 MATHEMATICS SUBJECT CLASSIFICATION
65xx

1. Introduction

Differential equations are ubiquitous in mathematical modelling of various phenomena. When we associate a mathematical model with physical phenomena, we often try to capture what is necessary, retaining the essential quantities and ignoring the negligible ones which involve small parameters. The model which contains small parameters is termed as a perturbed model, whereas the degenerate simplified model is called an unperturbed or reduced model. If a small parameter is affecting the highest differential coefficient, the problem is said to be singularly perturbed. The solution of these equations exhibits multiscale character since the corresponding degenerate system fails to satisfy the given boundary data. There are narrow regions across which solution changes rapidly and displays layer behaviour. Standard numerical methods on uniform meshes do not provide an efficient approximation to solutions in layer regions. Indeed, finding a consistent numerical approximation becomes more challenging for problems with discontinuous coefficients. The solution of the problem exhibits steep gradients alongside points of discontinuities and consequently leads to interior layers. Moreover, the prerequisite requirement of location and width of layers is too restrictive to be practical.

Efforts have been made by many researchers to provide consistent numerical approximations to singularly perturbed differential equations with discontinuous coefficients and source terms. In [30,31], a standard numerical method with piecewise linear interpolation

CONTACT Aditya Kaushik akaushik@dtu.ac.in, apkaushik@gmail.com Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India; University Institute of Engineering and Technology, Panjab University, Chandigarh 160014, India

© 2020 Informa UK Limited, trading as Taylor & Francis Group

on a Shishkin mesh is proposed to solve second and third-order ordinary differential equations. In [7], the authors solve a reaction–diffusion equation with a discontinuous diffusion coefficient. Whereas, in [27] the authors present an almost first-order uniformly convergent method for a coupled system of two reaction–diffusion equations. Later, in [25] the method was extended to n number of equations.

A discrete approximation of singularly perturbed parabolic partial differential equation with a discontinuous initial condition is proposed in [12]. The problem is bi-singular, where a classical singularity is twisted together with the singular nature of the problem. A numerical scheme is presented based on the fitted operator method defined over a uniform mesh. Indeed, it is an exact finite difference scheme for the error function associated with the discontinuity in the initial condition. However, there appears no further attempt towards developing fitted operator methods for parabolic reaction–diffusion problems until authors in [11] regularize the problem by reinstating the discontinuous initial condition. A linear parabolic problem with a turning point is also brought to attention [23]. The interior layer problem considered by the authors is of the following form

$$\begin{aligned} &-\varepsilon u_{xx} + a u_x + bu + cu_t = f, \quad (x,t) \in (0,1) \times (0,T], \quad b, c \geq 0 \\ &0 < \varepsilon \ll 1, \quad u(0,t),\ u(1,t),\ u(x,0) \text{ is specified.} \end{aligned} \tag{1}$$

A particular case, when the convective coefficient is discontinuous across a curve $\Gamma_1 := \{(d(t), t) \mid t \in [0,T],\ \ 0 < d(t) < 1\}$ and having a particular sign pattern $a(x) > 0$, $x < d(t)$ or $a(x) < 0$, $x > d(t)$ is examined in [14]. Whereas, in [24] the author presents a parameter uniform method over a specially designed Shishkin mesh. The mesh is obtained by mapping Γ_1 to the vertical line $x = d(0)$. The case when the initial condition $u(x,0)$ contains its interior layer is studied separately in [10]. However, the convective coefficient is assumed to be space independent, smooth, and of one sign. It is to note that the reduced initial condition (set $\varepsilon = 0$) is discontinuous at some point. Moreover, the discontinuity travels with the characteristic curve $\Gamma_2 := \{(d(t), t) \mid t \in [0,T], d'(t) = a(t),\ d(0) = d\}$ associated with the reduced hyperbolic equation $av_x + bv + cv_t = f$. In [19], a hybrid difference scheme is presented to solve the parabolic problem with a discontinuous convection coefficient. On the other hand, in [23], an interior layer appears in the solution of (1) since the convective coefficient of the problem contains an interior layer with a hyperbolic tangent profile. This problem appears to be a time-dependent version of the ordinary differential equation examined in [22]. Whereas, ref. [6] deals with a parabolic problem with a boundary turning point. In [8], an experimental technique is presented to analyse uniform convergence over piecewise uniform meshes. However, it was presented in context to a singularly perturbed elliptic equation when convective term degenerates on the boundary of the domain.

In [4,5], a two-parameter parabolic problem is studied. A hybrid monotone difference scheme is conferred using the method of averaging at the discontinuous points [5]. While [4] presents the case when convective term degenerates inside the spatial domain, and the source term has a discontinuity of the first kind on the degeneration line. A particular case, when $\mu = 0$ was studied in [9]. Whereas, in [2,3,15], a similar problem for $\mu = 1$ was studied. In [28], the authors present a parameter uniformly convergent method for a singularly perturbed parabolic system of equations with a discontinuous source term. In this work, the diffusion term in each equation is affected by a small positive parameter of different magnitude. Besides, attention has been paid to nonlinear problems with an interior layer

[16,20,26]. In [16], the authors consider a parabolic periodic boundary value problem and construct the interior layer type formal asymptotics. The existence and asymptotic stability of the solution is established using precise lower and upper solutions. The analysis presented is then used to construct an efficient numerical method for a slightly general nonlinear problem in [26].

The analysis of special methods for singularly perturbed parabolic partial functional differential equation with discontinuous data and degenerating convective terms has seen little developed in the literature. In [30], a finite difference method based on linear interpolation is used to solve an ordinary delay differential equation with discontinuous coefficient. Later, a hybrid initial value method is considered for the same problem and showed that the method gives better result. In this paper, we give attention to solve parabolic delay differential equation with discontinuous coefficients.

We organize the paper as follows: In Section 2, we formulate the problem. Section 3 describes the semidiscretization of the problem in the time direction and gives ε-uniform error estimate in t-direction. In Section 4, we define the piecewise uniform Shishkin mesh and discuss the fully discrete implicit scheme. Section 5 provides parameter-uniform error estimate for the finite difference scheme in the spatial direction. Section 6 presents some test examples and the main conclusion is given in the next section.

2. Boundary value problem

Let $D = S^- \cup S^+ = (0,1) \times (0,T] \cup (1,2) \times (0,T]$ and consider the nonhomogeneous initial-boundary-value problem

$$\begin{aligned}
&\varepsilon u_{xx}(x,t) + a(x)u_x(x,t) - b(x)u(x-1,t) - c(x)u(x,t) - u_t(x,t) = f(x,t) \quad \text{in } D\\
&u(x,t) = p_0(x) \quad \text{on } [0,2] \times \{t=0\}\\
&u(x,t) = p_1(x,t) \quad \text{in } [-1,0] \times [0,T], \quad u(x,t) = p_2(t) \quad \text{on } \{x=2\} \times [0,T]
\end{aligned} \tag{2}$$

where $\varepsilon \ll 1$ is a small positive parameter, b and c are sufficiently smooth functions such that $b(x) < 0$, $c(x) > 0$ and $b(x) + c(x) \geq 0$ for all $x \in [0,2]$. Moreover, we assume that

$$\begin{aligned}
&a(x) = \begin{cases} a_1(x) & \text{if } 0 \leq x \leq 1 \\ a_2(x) & \text{if } 1 < x \leq 2, \end{cases} \quad f(x,t) = \begin{cases} f_1(x,t) & \text{if } (x,t) \in \overline{S}^- \\ f_2(x,t) & \text{if } (x,t) \in \overline{S}^+ \end{cases}\\
&a_1(x) < -\gamma_1 < -2\gamma < 0, \quad a_2(x) > \gamma_2 > 2\gamma > 0\\
&|[a]| \leq C, \quad |[f]| \leq C
\end{aligned} \tag{3}$$

where $\gamma = \min\{\gamma_1, \gamma_2\}$ and compatibility of the boundary data. The solution of (2) satisfies $[u] = 0$ and $[u_x] = 0$ at $x = 1$. Here, $[u]$ denotes the jump of u defined at the point of discontinuity $x = 1$ as $[u](1,t) = u(1^+,t) - u(1^-,t)$ where $u(1^\pm,t) = \lim_{x \to 1 \pm 0} u(x,t)$. The functions p_0, p_1 and p_2 are Hölder continuous and the compatibility conditions hold at the corners of the domain. Let us rewrite (2) as

$$L_\varepsilon u(x,t) = \mathcal{F}(x,t)$$

where

$$L_\varepsilon u = \begin{cases} \varepsilon u_{xx}(x,t) + a_1(x)u_x(x,t) - c(x)u(x,t) - u_t(x,t) & \text{if } (x,t) \in S^- \\ \varepsilon u_{xx}(x,t) + a_2(x)u_x(x,t) & \\ \quad -b(x)u(x-1,t) - c(x)u(x,t) - u_t(x,t) & \text{if } (x,t) \in S^+ \end{cases}$$

and

$$\mathcal{F}(x,t) = \begin{cases} f_1(x,t) + b(x)p_1(x-1,t) & \text{if } (x,t) \in S^- \\ f_2(x,t) & \text{if } (x,t) \in S^+. \end{cases}$$

Under these assumptions, the solution of (2) exists and is unique [1,13]. The simultaneous presence of discontinuity and delay make the problem stiff. The solution $u(x,t)$ of (2) exhibits a strong interior layer and weak boundary layer in the neighbourhood of the point $x = 1$ and $x = 2$, respectively. Moreover, it is easy to follow that the differential operator L_ε satisfies the following minimum principle.

Lemma 2.1: *Suppose* $\mathcal{P} \in C^0(\bar{D}) \cap C^2(S^- \cup S^+)$ *satisfies* $\mathcal{P}(x,t) \geq 0$ *for all* $(x,t) \in \Gamma := \bar{D} \setminus D$ *and* $L_\varepsilon \mathcal{P}(x,t) \leq 0$ *for all* $(x,t) \in D$. *Then* $\mathcal{P}(x,t) \geq 0$ *for all* $(x,t) \in \bar{D}$.

Proof: Choose $(x^k, t^k) \in D$ such that $\mathcal{P}(x^k, t^k) = \min_{(x,t)\in\bar{D}} \mathcal{P}(x,t)$. Consequently,

$$\mathcal{P}_x(x^k, t^k) = 0, \quad \mathcal{P}_t(x^k, t^k) = 0 \quad \text{and} \quad \mathcal{P}_{xx}(x^k, t^k) > 0.$$

Suppose $\mathcal{P}(x^k, t^k) < 0$ and it follows that $(x^k, t^k) \notin \Gamma$. If $(x^k, t^k) \in S^- \cup S^+$, note that $L_\varepsilon \mathcal{P}(x^k, t^k) > 0$. A contradiction to the assumption and consequently the required result follows. ■

An important application of the minimum principle is establishing the boundedness of solution. As an immediate application we obtain

Lemma 2.2: *Let u be the solution of* (2). *Then*

$$\|u\|_{\infty,\bar{D}} \leq \|u\|_{\infty,\Gamma} + \frac{1}{\gamma}\|f\|_{\infty,\bar{D}}, \quad \gamma = \min\{\gamma_1, \gamma_2\}. \tag{4}$$

Proof: Consider

$$\psi_\pm = \begin{cases} \|u\|_{\infty,\Gamma} + \dfrac{x}{\gamma}\|f\|_{\infty,\bar{D}} \pm u & \text{if } x \leq 1 \\ \|u\|_{\infty,\Gamma} + \dfrac{(2-x)\|f\|_{\infty,\bar{D}}}{\gamma} \pm u & \text{if } x \geq 1. \end{cases}$$

For $(x,t) \in S^-$ it follows that

$$L_\varepsilon \psi_\pm(x,t) = \pm L_\varepsilon u + a_1(x)\frac{\|f\|}{\gamma} - c(x)\|u\| - c(x)\frac{x\|f\|}{\gamma} \leq 0$$

since $a_1(x) \leq 0$ and $c(x) > 0$. Similarly, for $(x,t) \in S^+$, it is easy to verify that $L_\varepsilon \psi_\pm(x,t) \leq 0$. The required result (4) now follows from Lemma 2.1. ■

In general, one can assume homogeneous boundary conditions $p_0 = p_1 = p_2 = 0$ by subtracting from u some suitable smooth function that satisfies the original boundary conditions [29]. Moreover, as proved in [13], it is an easy exercise to obtain the following estimate.

Lemma 2.3: *Let u be the solution of* (2). *Then*

$$\left|\frac{\partial^i u}{\partial t^i}(x,t)\right| \le C \quad \text{for all } (x,t) \in \bar{D} \text{ and } i = 0,1,2.$$

Proof: Follows from the mean value theorem and (4). ■

3. Time discretization

Let $\mathbb{T}_t^M = \{t_k = kT/M, k = 0,\dots,M\}$ be a uniform mesh in the time direction. Then, the backward Euler method leads to

$$\begin{aligned}&\varepsilon\Delta t U_{xx}(x,t_{k+1}) + a\Delta t U_x(x,t_{k+1}) - b\Delta t U(x-1,t_{k+1}) - \hat{c}U(x,t_{k+1})\\ &\quad = -U(x,t_k) + \Delta t f(x,t_{k+1}), \quad x \in [0,2] \text{ and } k = 0,1,\dots,M-1\end{aligned} \tag{5}$$

such that

$$\begin{aligned}&U(x,0) = p_0(x), \quad 0 \le x \le 2,\\ &U(x,t_{k+1}) = p_1(x,t_{k+1}), \quad -1 \le x \le 0,\ 0 \le k \le M-1,\\ &U(2,t_{k+1}) = p_2(t_{k+1}), \quad 0 \le k \le M-1\end{aligned} \tag{6}$$

where $\hat{c}(x) = (c(x)\Delta t + 1)$. Let us rewrite (5) as

$$\mathfrak{L}^M U(x,t_{k+1}) = \mathcal{F}(x,t_{k+1}) \tag{7}$$

where

$$\mathfrak{L}^M U = \begin{cases} \varepsilon\Delta t U_{xx}(x,t_{k+1}) + a(x)\Delta t U_x(x,t_{k+1}) - \hat{c}(x)U(x,t_{k+1}) & \text{if } x \in [0,1]\\ \varepsilon\Delta t U_{xx}(x,t_{k+1}) + a(x)\Delta t U_x(x,t_{k+1}) - b(x)\Delta t U(x-1,t_{k+1}) & \\ \quad -\hat{c}(x)U(x,t_{k+1}) & \text{if } x \in (1,2)\end{cases} \tag{8}$$

and

$$\mathcal{F}(x,t_{k+1}) = \begin{cases} \Delta t f(x,t_{k+1}) + \Delta t b(x) p_1(x-1,t_{k+1}) - U(x,t_k) & \text{if } x \in (0,1]\\ \Delta t f(x,t_{k+1}) - U(x,t_k) & \text{if } x \in (1,2].\end{cases} \tag{9}$$

The operator $\mathfrak{L}^M$ satisfies the following minimum principle.

Lemma 3.1: *Let $\phi(x,t_{k+1})$ be a smooth function such that $\phi(x,t_{k+1}) \ge 0$ for $x = 0, 2$ and $\mathfrak{L}^M\phi(x,t_{k+1}) \le 0$ for all $x \in (0,2)$. Then $\phi(x,t_{k+1}) \ge 0$ for all $x \in [0,2]$.*

Proof: Choose (x^o, t_s^o) such that $\phi(x,t_{k+1})$ attains its minimum at (x^o,t_s^o). Then

$$\phi_x(x^o,t_s^o) = \phi_t(x^o,t_s^o) = 0 \quad \text{and} \quad \phi_{xx}(x^o,t_s^o) > 0. \tag{10}$$

Suppose $\phi(x^o,t_s^o) < 0$ and it follows that $(x^o,t_s^o) \notin \Gamma$ since $\phi(x,t_{k+1}) \ge 0$ for $x = 0, 2$.

Case 1: If $x^o \in S^-$

$$\begin{aligned}\mathfrak{L}^M\phi(x^o,t_s^o) &= \varepsilon\Delta t\phi_{xx}(x^o,t_s^o) + a_1(x^o)\Delta t\phi_x(x^o,t_s^o) - \hat{c}(x^o)\phi(x^o,t_s^o)\\ &> 0 \text{ from (3) and (10).}\end{aligned}$$

Case 2: If $x^o \in S^+$

$$\begin{aligned}\mathfrak{L}^M\phi(x^o,t_s^o) &= \varepsilon\Delta t\phi_{xx}(x^o,t_s^o) + a_2(x^o)\Delta t\phi_x(x^o,t_s^o)\\ &\quad - b(x^o)\Delta t\phi(x^o-1,t_s^o) - \hat{c}(x^o)\phi(x^o,t_s^o)\\ &= \varepsilon\Delta t\phi_{xx}(x^o,t_s^o) + a_2(x^o)\Delta t\phi_x(x^o,t_s^o)\\ &\quad - b(x^o)\Delta t(\phi(x^o-1,t_s^o) - \phi(x^o,t_s^o))\\ &\quad - b(x^o)\Delta t\phi(x^o,t_s^o) - c(x^o)\Delta t\phi(x^o,t_s^o) - \phi(x^o,t_s^o)\\ &= \varepsilon\Delta t\phi_{xx}(x^o,t_s^o) + a_2(x^o)\Delta t\phi_x(x^o,t_s^o)\\ &\quad - b(x^o)\Delta t(\phi(x^o-1,t_s^o) - \phi(x^o,t_s^o))\\ &\quad - \Delta t\phi(x^o,t_s^o)(b(x^o)+c(x^o)) - \phi(x^o,t_s^o)\\ &> 0 \text{ from (3) and (10).}\end{aligned}$$

It contradicts the assumption and hence, the required result follows. ■

The operator $\mathfrak{L}^M$ is inverse monotone and $\|(\mathfrak{L}^M)^{-1}\| \le C$. Consequently, the stability of the scheme is immediate.

Lemma 3.2: *Let $\hat{e}_{k+1} = \hat{U}(x,t_{k+1}) - u(x,t_{k+1})$ be the local truncation error at $k+1$th time step. Then $\|\hat{e}_{k+1}\|_\infty \le C(\Delta t)^2$ for some constant C.*

Moreover, if $E_k = u(x,t_k) - U(x,t_k)$ denotes the global error in the time direction. Then, it follows that

$$\|E_{k+1}\|_\infty = \left\|\sum_{i=1}^{k}\hat{e}_i\right\|_\infty \le \|\hat{e}_1\|_\infty + \|\hat{e}_2\|_\infty + \|\hat{e}_3\|_\infty + \cdots + \|\hat{e}_k\|_\infty \le C\Delta t. \tag{11}$$

This in turn ensures the uniform convergence of time semidiscretization process. Next, we obtain a priori estimate on the solution of semidiscretized problem (5).

Lemma 3.3: *Let $U(x,t_{k+1})$ be the solution of* (5)*. Then*

$$|U(x,t_{k+1})| \le \max\left\{|U(0,t_{k+1})|, \frac{||\mathcal{F}||}{\gamma}, |U(2,t_{k+1})|\right\} \quad \text{for all } x \in [0,2].$$

Proof: Consider

$$\psi^{\pm}(x,t_{k+1}) = \max\left\{|U(0,t_{k+1})|, \frac{||\mathcal{F}||}{\gamma}, |U(2,t_{k+1})|\right\} \pm U(x,t_{k+1}).$$

Then, $\psi^{\pm}(0,t_{k+1}) \geq 0$ and $\psi^{\pm}(2,t_{k+1}) \geq 0$. Moreover, for $x \in [0,1]$

$$\begin{aligned}\mathfrak{L}^M\psi^{\pm}(x,t_{k+1}) &= -\hat{c}(x)\max\left\{|U(0,t_{k+1})|, \frac{||\mathcal{F}||}{\gamma}, |U(2,t_{k+1})|\right\} \pm \mathfrak{L}^M U(x,t_{k+1})\\ &\leq 0\end{aligned}$$

since $\hat{c}(x) = c(x)\Delta t + 1 \geq 0$. Similarly, for $x \in [1,2]$ we compute

$$\begin{aligned}\mathfrak{L}^M\psi^{\pm}(x,t_{k+1}) &= -(b(x)\Delta t + \hat{c}(x))\max\left\{|U(0,t_{k+1})|, \frac{||\mathcal{F}||}{\gamma}, |U(2,t_{k+1})|\right\}\\ &\pm \mathfrak{L}^M U(x,t_{k+1}) \leq 0\end{aligned}$$

since $b(x) + c(x) \geq 0$. Consequently, from Lemma 3.1 the required result follows. ■

The solution $U(x,t_{k+1})$ of the semidiscretized problem (5) is known to admits a decomposition into smooth and a singular component [17]. We write

$$U(x,t_{k+1}) := V(x,t_{k+1}) + W(x,t_{k+1}).$$

The smooth component $V(x,t_{k+1})$ is the solution of

$$\begin{aligned}&\mathfrak{L}^M V(x,t_{k+1}) = \mathcal{F}(x,t_{k+1}), \quad x \in (0,2)\\ &V(0,t_{k+1}) = V_0(0,t_{k+1}),\\ &V(1^-,t_{k+1}) = -(\hat{c}(1))^{-1}(\Delta t f(1,t_{k+1}) - U(1,t_k) + \Delta t b(1) p_1(0,t_{k+1})),\\ &V(1^+,t_{k+1}) = -(\hat{c}(1))^{-1}(\Delta t f(1,t_{k+1}) - U(1,t_k) + \Delta t b(1) V_0(0,t_{k+1})),\\ &V(2,t_{k+1}) = V_0(2,t_{k+1}),\end{aligned} \tag{12}$$

where $V_0(x,t_{k+1})$ satisfies the corresponding degenerate problem. Also, the singular component $W(x,t_{k+1})$ satisfies a homogeneous problem which reads

$$\begin{aligned}&\mathfrak{L}^M W(x,t_{k+1}) = 0, \quad x \in (0,2)\\ &W(x,t_{k+1}) = U(x,t_{k+1}) - V(x,t_{k+1}), \quad x = 0,2\\ &W(1^+,t_{k+1}) - W(1^-,t_{k+1}) = V(1^-,t_{k+1}) - V(1^+,t_{k+1}),\\ &W_x(1^+,t_{k+1}) - W_x(1^-,t_{k+1}) = V_x(1^-,t_{k+1}) - V_x(1^+,t_{k+1}).\end{aligned} \tag{13}$$

The next two lemmas provide bounds on the derivative of $V(x,t_{k+1})$ and $W(x,t_{k+1})$ with respect to x.

Lemma 3.4: *Let $V(x, t_{k+1})$ and $W(x, t_{k+1})$ be the solutions of* (12) *and* (13)*, respectively. Then, for $k = 0, 1, 2, 3$*

$$\left|\frac{\mathrm{d}^k V(x, t_{k+1})}{\mathrm{d}x^k}\right| \leq C(1 + \varepsilon^{2-k})$$

for $x \in (0, 1) \cup (1, 2)$ and

$$\left|\frac{\mathrm{d}^k W(x, t_{k+1})}{\mathrm{d}x^k}\right| \leq \begin{cases} \varepsilon^{-k} \exp\left(\dfrac{-\gamma(1-x)}{\varepsilon}\right) & \text{for } x \in (0, 1) \\ \varepsilon^{-k} \exp\left(\dfrac{-\gamma(x-1)}{\varepsilon}\right) + \varepsilon^{-(k+1)} \exp\left(\frac{-\gamma(2-x)}{\varepsilon}\right) & \text{for } x \in (1, 2). \end{cases}$$

Proof: The proof follows from [30]. ■

4. Spatial discretization

The solution of the problem exhibits strong interior layer across discontinuity $x = 1$ and a weak boundary layer at $x = 2$. Therefore, we construct a piecewise-uniform mesh $\bar{\mathbb{D}}_x^N$ which condenses around $x = 1$ and $x = 2$. We write

$$[0, 2] = [0, 1 - \tau_1] \cup [1 - \tau_1, 1] \cup [1, 1 + \tau_2] \cup [1 + \tau_2, 2 - \tau_2] \cup [2 - \tau_2, 2]$$

where $\tau_1 = \min\{0.5, \frac{2\varepsilon \ln N}{\gamma}\}$ and $\tau_2 = \min\{0.25, \frac{2\varepsilon \ln N}{\gamma}\}$ are the mesh transition parameters. We place $N/4$ mesh points each in intervals $[0, 1 - \tau_1]$, $[1 - \tau_1, 1]$, $[1 + \tau_2, 2 - \tau_2]$ and $N/8$ mesh points in intervals $[1, 1 + \tau_2]$ and $[2 - \tau_2, 2]$. Consequently, we obtain

$$\bar{\mathbb{D}}_x^N = \{x_i\}_0^N = \begin{cases} x_i = 0 & \text{for } i = 0, \\ x_i = x_{i-1} + h_i & \text{for } i = 2, \ldots, N \end{cases}$$

where

$$h_i = \begin{cases} \dfrac{4}{N}(1 - \tau_1) & \text{for } i = 2 : N/4, \\ \dfrac{4}{N}\tau_1 & \text{for } i = N/4 + 1 : N/2, \\ \dfrac{8}{N}\tau_2 & \text{for } i = N/2 + 1 : 5N/8, \\ \dfrac{4}{N}(1 - 2\tau_2) & \text{for } i = 5N/8 + 1 : 7N/8, \\ \dfrac{8}{N}\tau_2 & \text{for } i = 7N/8 + 1 : N. \end{cases}$$

To discretize the differential operator in (5), we define the finite difference operators on the piecewise uniform mesh $\bar{\mathbb{D}}_x^N$ as

$$D_x^+ U_{i,k+1} = \frac{U_{i+1,k+1} - U_{i,k+1}}{h_{i+1}}, \quad D_x^- U_{i,k+1} = \frac{U_{i,k+1} - U_{i-1,k+1}}{h_i} \text{ and}$$

$$\delta^2 U_{i,k+1} = D_x^+ D_x^- U_{i,k+1} = \frac{2(D_x^+ U_{i,k+1} - D_x^- U_{i,k+1})}{h_i + h_{i+1}}.$$

The discrete problem on $\bar{\mathbb{D}}_x^N \times \mathbb{T}_t{}^M$ thus reads

$$\mathfrak{L}_\varepsilon^{N,M} U_{i,k+1} = \mathcal{F}_{i,k+1} \tag{14}$$

where

$$\mathfrak{L}_\varepsilon^{N,M} U(x_i, t_{k+1}) = \begin{cases} \varepsilon \Delta t D_x^+ D_x^- U_{i,k+1} & \\ \quad + a_i \Delta t D_x^- U_{i,k+1} - \hat{c}_i U_{i,k+1} & \text{for } \le i \le N/2 - 1 \\ \varepsilon \Delta t D_x^+ D_x^- U_{i,k+1} + a_i \Delta t D_x^+ U_{i,k+1} & \\ \quad - b_i \Delta t U_{i-N/2,k+1} - \hat{c}_i U_{i,k+1} & \text{for } N/2 + 1 \le i \le N - 1 \end{cases}$$

and

$$\mathcal{F}_{i,k+1} = \begin{cases} \Delta t f_{i,k+1} - U_{i,k} + \Delta t b_i p_1(i - N/2, k+1), & \text{for } 1 \le i \le N/2 - 1, \\ \Delta t f_{i,k+1} - U_{i,k}, & \text{for } N/2 + 1 \le i \le N - 1. \end{cases}$$

Moreover, for $i = N/2$

$$D_x^+ U_{N/2,k+1} = D_x^- U_{N/2,k+1}$$

with

$$\begin{aligned} &U_{i,0} = p_0(x_i), \quad \text{for } i \in \{0, \dots, N\} \\ &U_{N,k+1} = p_2(t_{k+1}), \quad \text{for } k \in \{0, 1, 2, \dots, M-1\} \\ &U_{i,k+1} = p_1(i, k+1), \quad \text{for } k \in \{0, 1, 2, \dots, M-1\} \text{ and } i \in \{-N/2, \dots 0\}. \end{aligned}$$

It is easy to follow from [13] that

Lemma 4.1: *Let $\phi_{i,k+1}$ be a mesh function so that $\phi_{i,k+1} \ge 0$ for $i = \{0, N\}$, $\mathfrak{L}_\varepsilon^{N,M} \phi_{i,k+1} \le 0$ for all $i \in \{1, \dots, N\} \backslash \{N/2\}$ and $D_x^+ \phi_{N/2,k+1} - D_x^- \phi_{N/2,k+1} \ge 0$. Then $\phi_{i,k+1} \ge 0$ for all $i = 0, 1, \dots, N$.*

Proof: Choose $i^* \in \{0, 1, \dots, N\} \backslash \{N/2\}$ such that $\phi_{i^*,k+1} = \min_{\bar{\mathbb{D}}_x^N \times \mathbb{T}_t{}^M} \phi_{i,k+1}$. Suppose $\phi_{i^*,k+1} < 0$ and it follows that $i^* \notin \{0, N\}$. For $i^* \in \{1, 2, \dots, N/2 - 1\}$

$$\begin{aligned} \mathfrak{L}_\varepsilon^{N,M} \phi_{i^*,k+1} &= \varepsilon \Delta t D_x^+ D_x^- \phi_{i^*,k+1} + a_i \Delta t D_x^- \phi_{i^*,k+1} - c_i \phi_{i^*,k+1} \\ &= \frac{2\varepsilon \Delta t}{\hat{h}_i} \left\{ \frac{\phi_{i^*+1,k+1} - \phi_{i^*,k+1}}{h_{i+1}} - \frac{\phi_{i^*,k+1} - \phi_{i^*-1,k+1}}{h_i} \right\} \\ &\quad + a_i \Delta t \left\{ \frac{\phi_{i^*,k+1} - \phi_{i^*-1,k+1}}{h_i} \right\} - c_i \phi_{i^*,k+1} \\ &\ge 0. \end{aligned}$$

Also, $\mathfrak{L}_\varepsilon^{N,M} \phi_{i^*,k+1} \ge 0$ for $i^* \in \{N/2 + 1, \dots, N - 1\}$. A contradiction to the assumption and consequently the required result follows. ■

As an immediate consequence of this Lemma we obtain the following estimate.

Lemma 4.2: *Let $\phi_{i,k+1}$ be the numerical solution of (14). Then*

$$|\phi_{i,k+1}| \leq \max\{|\phi_{0,k+1}|, |\phi_{N,k+1}|, \|\mathfrak{L}_{\varepsilon}^{N,M}\phi\|\}, \quad \forall\, i = 0:N,\ k = 0:M-1.$$

Proof: Let $\psi_{i,k+1}^{\pm} = \max\{|\phi_{0,k+1}|, |\phi_{N,k+1}|, \|\mathfrak{L}_{\varepsilon}^{N,M}\phi\|\} \pm \phi_{i,k+1}$. Clearly, $\psi_{i,k+1}^{\pm} \geq 0$ for $i = 0, N$ and $\mathfrak{L}_{\varepsilon}^{N,M}\psi_{i,k+1}^{\pm} \leq 0$ for $i \in \{1, \ldots, N-1\}\backslash N/2$. Moreover, $(D_x^+ - D_x^-)\psi_{i,k+1}^{\pm} = 0$ for $i = N/2$. The required result thus follows from Lemma 4.1. ■

5. Error estimates

To obtain parameter uniform error bounds we decompose the solution $U_{i,k+1}$ into regular and singular components. We write $U_{i,k+1} = V_{i,k+1} + W_{i,k+1}$, where V and W satisfy

$$\begin{aligned}
&\mathfrak{L}_{\varepsilon}^{N,M} V_{i,k+1} = \mathcal{F}_{i,k+1} \quad \text{for } i \in \{1, \ldots, N-1\} \setminus N/2,\\
&V_{0,k+1} = V(0, t_{k+1}),\\
&V_{N/2-1,k+1} = V(1^-, t_{k+1}),\\
&V_{N/2+1,k+1} = V(1^+, t_{k+1}),\\
&V_{N,k+1} = V(2, t_{k+1})
\end{aligned} \tag{15}$$

and

$$\begin{aligned}
&\mathfrak{L}_{\varepsilon}^{N,M} W_{i,k+1} = 0, \quad \text{for } i \in \{1, \ldots, N-1\} \setminus N/2,\\
&W_{0,k+1} = W(0, t_{k+1}),\\
&W_{N,k+1} = W(2, t_{k+1}),\\
&V_{N/2+1,k+1} + W_{N/2+1,k+1} = V_{N/2-1,k+1} + W_{N/2-1,k+1},\\
&D_x^- V_{N/2,k+1} + D_x^- W_{N/2,k+1} = D_x^+ V_{N/2,k+1} + D_x^+ W_{N/2,k+1}.
\end{aligned} \tag{16}$$

Moreover, the error $e_{i,k+1}$ is defined as

$$\begin{aligned}
e_{i,k+1} &= U(x_i, t_{k+1}) - U_{i,k+1}\\
&= (V(x_i, t_{k+1}) - V_{i,k+1}) + (W(x_i, t_{k+1}) - W_{i,k+1}).
\end{aligned}$$

Theorem 5.1: *Let $V(x_i, t_{k+1})$ and $V_{i,k+1}$ be the solution of* (12) *and* (15)*, respectively. Then*

$$|\mathfrak{L}_{\varepsilon}^{N,M}(V(x_i, t_{k+1}) - V_{i,k+1})| \leq CN^{-1} \quad \text{for } i \in \{0, \ldots, N\}\backslash\{N/2\}.$$

Moreover, if $W(x_i, t_{k+1})$ and $W_{i,k+1}$ be the solution of (13) *and* (16)*, respectively. Then*

$$|\mathfrak{L}_{\varepsilon}^{N,M}(W(x_i, t_{k+1}) - W_{i,k+1})| \leq CN^{-1}(\ln N)^2 \quad \text{for } i \in \{0, \ldots, N\}\backslash\{N/2\}.$$

Proof: The proof imitates [13,18]. ■

Next, we obtain some priori estimates at $(x_{N/2}, t_{k+1})$. To that end we consider

$$\begin{aligned}
|(D_x^+ - D_x^-)e_{N/2,k+1}| &= |(D_x^+ - D_x^-)(U(x_{N/2}, t_{k+1}) - U_{N/2,k+1})| \\
&= |(D_x^+ - D_x^-)(U(x_{N/2}, t_{k+1}))| \\
&\le \left| \left(D_x^+ - \frac{\mathrm{d}}{\mathrm{d}x} \right) U(x_{N/2}, t_{k+1}) \right| + \left| \left(D_x^- - \frac{\mathrm{d}}{\mathrm{d}x} \right) U(x_{N/2}, t_{k+1}) \right| \\
&\le \frac{1}{2} h_{N/2}^+ \max_{x_i \in (1,2)} \left| \frac{\mathrm{d}^2 U(x_i, t_{k+1})}{\mathrm{d}x^2} \right| + \frac{1}{2} h_{N/2}^- \max_{x_i \in (0,1)} \left| \frac{\mathrm{d}^2 U(x_i, t_{k+1})}{\mathrm{d}x^2} \right| \\
&\le Ch^* \max_{x_i \in (0,1) \cup (1,2)} \left| \frac{\mathrm{d}^2 U(x_i, t_{k+1})}{\mathrm{d}x^2} \right| \\
&\le \frac{Ch^*}{\varepsilon^2} \qquad (17)
\end{aligned}$$

since $(D_x^+ - D_x^-)U_{N/2,k+1} = 0$ and $h^* = h_{N/2}^- = h_{N/2}^+$.

If we consider

$$\phi_{i,k+1} = \begin{cases} \dfrac{\prod_{k=1}^{i} \left(1 + \frac{\gamma}{\varepsilon} h_k\right)}{\prod_{k=1}^{N/2} \left(1 + \frac{\gamma}{\varepsilon} h_k\right)} & \text{for } 0 \le i \le N/2, \\ \dfrac{\prod_{k=i}^{N-1} \left(1 + \frac{\gamma}{\varepsilon} h_{k+1}\right)}{\prod_{k=N/2}^{N-1} \left(1 + \frac{\gamma}{\varepsilon} h_{k+1}\right)} & \text{for } N/2 \le i \le N. \end{cases}$$

then $0 \le \phi_{i,k+1} \le 1$ for $0 \le i \le N$. Moreover, for $i = 0, 1, \ldots, N/2$

$$\begin{aligned}
D_x^+ \phi_{i,k+1} &= \frac{\gamma}{\varepsilon} \phi_{i,k+1}, \\
D_x^- \phi_{i,k+1} &= \frac{1}{h_i} \frac{\prod_{k=1}^{i} \left(1 + \frac{\gamma}{\varepsilon} h_k\right)}{\prod_{k=1}^{N/2} \left(1 + \frac{\gamma}{\varepsilon} h_k\right)} \left(1 - \frac{1}{1 + \frac{\gamma}{\varepsilon} h_i} \right) = \frac{\gamma}{\varepsilon} \frac{1}{(1 + \frac{\gamma}{\varepsilon} h_i)} \phi_{i,k+1}
\end{aligned}$$

and

$$\begin{aligned}
\delta^2 \phi_{i,k+1} &= \frac{2}{h_i + h_{i+1}} (D_x^+ - D_x^-) \phi_{i,k+1} \\
&= \left(\frac{\gamma}{\varepsilon} \right)^2 \frac{2h_i}{h_i + h_{i+1}} \left(\frac{1}{1 + \frac{\gamma}{\varepsilon} h_i} \right) \phi_{i,k+1} \\
&\le 2 \left(\frac{\gamma}{\varepsilon} \right)^2 \phi_{i,k+1}.
\end{aligned}$$

Also, for $i = N/2, \ldots, N$

$$\begin{aligned}
D_x^+ \phi_{i,k+1} &= \frac{-\gamma}{\varepsilon} \left(\frac{\phi_{i,k+1}}{1 + \frac{\gamma}{\varepsilon} h_{i+1}} \right), \\
D_x^- \phi_{i,k+1} &= \frac{-\gamma}{\varepsilon} \phi_{i,k+1}, \quad \delta^2 \phi_{i,k+1} \le 2 \left(\frac{\gamma}{\varepsilon} \right)^2 \phi_{i,k+1}
\end{aligned}$$

and at $i = \frac{N}{2}$

$$\begin{aligned}(D_x^+ - D_x^-)\phi_{i,k+1} &= -\frac{\gamma}{\varepsilon}\left(\frac{1}{1+\frac{\gamma}{\varepsilon}h_{N/2}^+} + \frac{1}{1+\frac{\gamma}{\varepsilon}h_{N/2}^-}\right)\phi_{N/2,k+1} \\ &= \frac{-\gamma}{\varepsilon}\left(\frac{2}{1+\frac{\gamma}{\varepsilon}h_{N/2}^*}\right)\phi_{N/2,k+1} \\ &\leq -\frac{C}{\varepsilon}.\end{aligned}$$

Then, for $1 \leq i \leq N/2 - 1$

$$\begin{aligned}\mathfrak{L}_\varepsilon^{N,M}\phi_{i,k+1} &= \varepsilon\Delta t\delta^2\phi_{i,k+1} + a_i\Delta t D_x^-\phi_{i,k+1} - \hat{c}_i\phi_{i,k+1} \\ &\leq \left(\frac{2\Delta t\gamma^2}{\varepsilon} + \frac{\gamma}{\varepsilon}\frac{a_i\Delta t}{(1+\frac{\gamma}{\varepsilon}h_i)} - \hat{c}_i\right)\phi_{i,k+1}. \qquad (18)\end{aligned}$$

Similarly, for $N/2 + 1 \leq i \leq N - 1$

$$\begin{aligned}\mathfrak{L}_\varepsilon^{N,M}\phi_{i,k+1} &= \varepsilon\Delta t\delta^2\phi_{i,k+1} + a_i\Delta t D_x^+\phi_{i,k+1} - b_i\Delta t\phi_{i-N/2,k+1} - \hat{c}_i\phi_{i,k+1} \\ &\leq \left(\frac{2\Delta t\gamma^2}{\varepsilon} - \frac{\gamma a_i\Delta t}{\varepsilon(1+\frac{\gamma}{\varepsilon}h_{i+1})} - \hat{c}_i\right)\phi_{i,k+1} - b_i\Delta t\phi_{i-N/2,k+1}.\end{aligned}$$

Theorem 5.2: *Let $U(x_i, t_{k+1})$ be the solution of* (5) *and $U_{i,k+1}$ be the solution of* (14). *Then at $(k+1)^{st}$ time step*

$$|U(x_i, t_{k+1}) - U_{i,k+1}| \leq CN^{-1}\ln^2 N, \quad \text{for } 0 \leq i \leq N.$$

Proof: Consider the barrier function

$$\psi_{i,k+1}^{\pm} = C_1N^{-1}\ln^2 N + C_2\frac{\gamma}{\varepsilon}h^*\phi_{i,k+1} \pm e_{i,k+1} \quad \text{for } i = 0, 1, \ldots, N$$

where C_1 and C_2 are constants. Using (18), the assumption $a_i < -2\gamma$ and Theorem 5.1 to obtain

$$\begin{aligned}\mathfrak{L}_\varepsilon^{N,M}\psi_{i,k+1}^{\pm} &= -C_1\hat{c}_iN^{-1}\ln^2 N + C_2\frac{\gamma}{\varepsilon}h^*\left(\frac{2\Delta t\gamma^2}{\varepsilon} + a_i\Delta t\frac{\gamma}{\varepsilon} - \hat{c}_i\right) \pm \mathfrak{L}_\varepsilon^{N,M}e_{i,k+1} \\ &\leq -C_1\hat{c}_iN^{-1}\ln^2 N + C_2\left(\frac{\gamma}{\varepsilon}\right)^2 h^*\Delta t\left(2\gamma + a_i - \hat{c}_i\frac{\varepsilon}{\gamma}\Delta t\right) \pm CN^{-1}\ln^2 N \\ &\leq 0\end{aligned}$$

for $i = 0, \ldots, N/2 - 1$. Similarly, for $i = N/2, \ldots, N$ we obtain

$$\begin{aligned}\mathfrak{L}_\varepsilon^{N,M}\psi_{i,k+1}^{\pm} &= C_1(-\hat{c}_i - b_i\Delta t)N^{-1}\ln^2 N + C_2\left(\frac{\gamma}{\varepsilon}\right)h^*\mathfrak{L}_\varepsilon^{N,M}\phi_{i,k+1} \pm \mathfrak{L}_\varepsilon^{N,M}e_{i,k+1}\\ &= C_1(-\hat{c}_i - b_i\Delta t)N^{-1}\ln^2 N + C_2\left(\frac{\gamma}{\varepsilon}\right)h^*\frac{\gamma}{\varepsilon}\Delta t\left(2\gamma - a_i - \hat{c}_i\frac{\varepsilon\Delta t}{\gamma}\right)\\ &\quad - C_2\left(\frac{\gamma}{\varepsilon}\right)h^* b_i\Delta t\phi_{i-N/2,k+1} \pm CN^{-1}\ln^2 N\\ &\le 0.\end{aligned}$$

Also, if $i = N/2$

$$\begin{aligned}(D_x^+ - D_x^-)\psi_{N/2,k+1}^{\pm} &= (D_x^+ - D_x^-)\left(C_1N^{-1}\ln^2 N + C_2\left(\frac{\gamma}{\varepsilon}\right)h^*\phi_{N/2,k+1} \pm e_{N/2,k+1}\right)\\ &\le C_2\left(\frac{\gamma}{\varepsilon}\right)h^*\left(\frac{-C}{\varepsilon}\right) + \frac{Ch^*}{\varepsilon^2}\\ &\le 0\end{aligned}$$

for some suitable constant C_2. Moreover, $\psi_{0,k+1}^{\pm} \ge 0$, $\psi_{N,k+1}^{\pm} \ge 0$ and $0 \le \phi_{i,k+1} \le 1$. Consequently, the required result follows from Lemma 4.1. ■

Next we combine (11) and Theorem 5.2 to obtain the main result.

Theorem 5.3: *Let u and $U_{i,k+1}$ are solution of continuous problem* (2) *and discrete problem* (14)*, respectively. Then*

$$|u(x_i, t_{k+1}) - U_{i,k+1}| \le C(\Delta t + (N^{-1}\ln^2 N))$$

for $0 \le i \le N$ and $0 \le k \le M$.

6. Numerical illustration

In this section, we examine the performance of the proposed method and numerically verify the theoretical estimates. We consider two test problems for numerical computations. The exact solution of the problem is not available for comparison. Therefore, we estimate the error using the double mesh principle [17]. The maximum point-wise error ($E_\varepsilon^{N,\Delta t}$) and order of convergence ($R_\varepsilon^{N,\Delta t}$) is calculated using

$$E_\varepsilon^{N,\Delta t} := \max|U^{N,\Delta t}(x_i, t_{k+1}) - \tilde{U}^{2N,\Delta t/2}(x_i, t_{k+1})| \quad \text{and } R_\varepsilon^{N,\Delta t} := \log_2\left(\frac{E_\varepsilon^{N,\Delta t}}{E_\varepsilon^{2N,\Delta t/2}}\right)$$

where $U^{N,\Delta t}(x_i, t_{k+1})$ and $\tilde{U}^{2N,\Delta t/2}(x_i, t_k)$ are the numerical solutions obtained on $\bar{\mathbb{D}}_x^N \times \mathbb{T}_t^M$ and $\bar{\mathbb{D}}_x^{2N} \times \mathbb{T}_t^{2M}$, respectively.

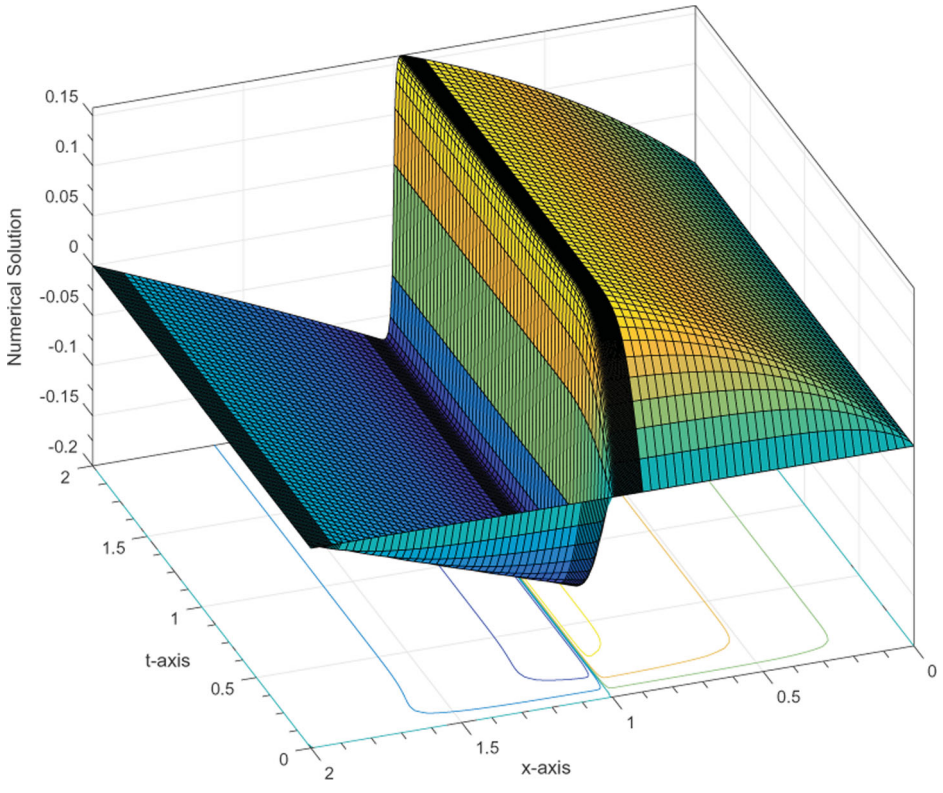


Figure 1. Numerical solution of Example 6.1 for $\varepsilon = 2^{-6}$ when $M = N = 128$.

Example 6.1: Consider the following singularly perturbed problem [13]:

$$\begin{aligned}
&\varepsilon u_{xx} + a(x)u_x + u(x-1,t) - 3u(x,t) - u_t = f(x,t), \quad (x,t) \in (0,2) \times (0,2],\\
&u(x,0) = 0, \quad x \in [0,2],\\
&u(x,t) = 0, \quad (x,t) \in [-1,0] \times [0,2],\\
&u(2,t) = 0, \quad t \in [0,2]
\end{aligned}$$

where

$$a(x) = \begin{cases} -(4+x), & x \in [0,1],\\ (3+x^2), & x \in (1,2] \end{cases} \quad \text{and} \quad f(x,t) = \begin{cases} -1, & (x,t) \in [0,1] \times [0,2],\\ 1, & (x,t) \in (1,2] \times [0,2]. \end{cases}$$

In case, perturbation parameter tends to zero, the solution of the problem exhibit turning point behaviour (Figures 1–4). Maximum absolute error and order of convergence for Examples 6.1 and 6.2 are tabulated for different values of ϵ, M, and N in Tables 1–4. Moreover, log–log plot of the maximum absolute error can be had from Figures 5 and 6. It is evident from it that the errors decreases monotonically as N increases. Numerical solution for Examples 6.1 and 6.2 are plotted in Figures 1 and 3, respectively. Also, the numerical solutions at final time step ($t = 2$) for different values of ϵ are displayed in Figures 2 and 4.

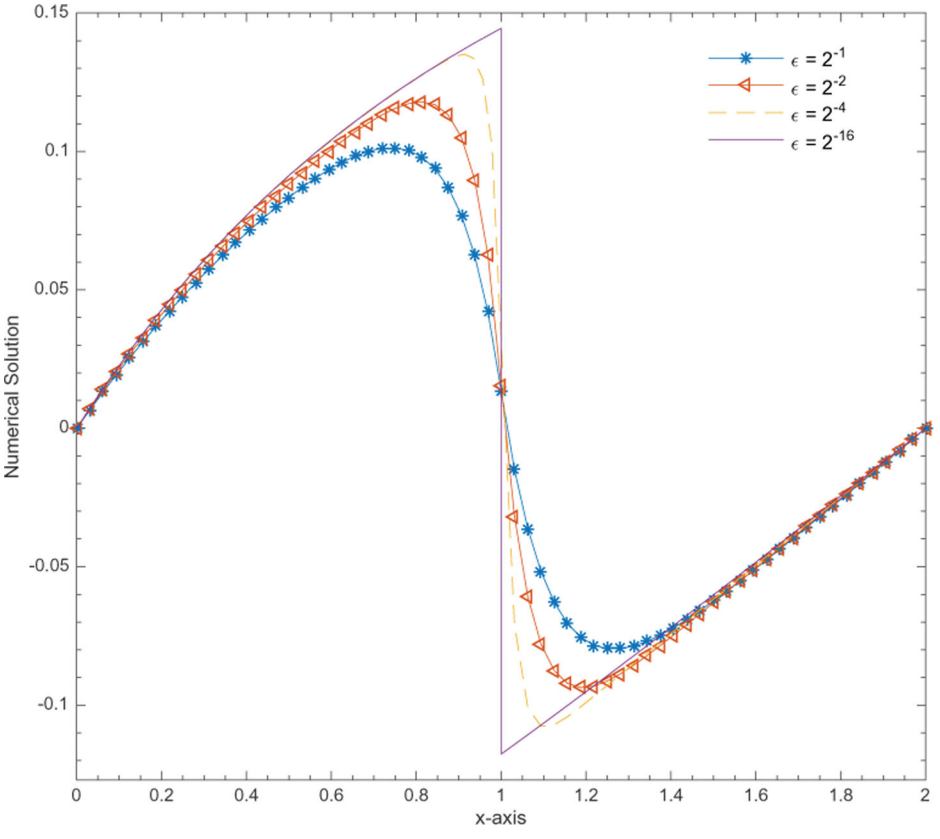


Figure 2. Numerical solutions of Example 6.1 at $t = 2$ for different values of ϵ when $N = 64$.

Table 1. Maximum absolute error and order of convergence for Example 6.1 for different values of ϵ, M, and N when $M = N$.

N	$\varepsilon = 2^{-0}$	2^{-1}	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}
32	0.007388	0.009292	0.013393	0.017789	0.019344	0.019860	0.019962
	0.9895	0.8712	0.8571	0.8718	0.8797	0.8639	0.8490
64	0.003721	0.005080	0.007394	0.009721	0.010513	0.010912	0.011082
	0.9834	0.9274	0.9196	0.9107	0.9253	0.8718	0.8571
128	0.001882	0.002671	0.003909	0.005171	0.005536	0.005963	0.006118
	0.9878	0.9423	0.9418	0.9263	0.9106	0.8876	0.8650
256	0.000949	0.001390	0.002035	0.002721	0.002945	0.003223	0.003359
	0.9924	0.9571	0.9588	0.9332	0.9260	0.9026	0.8809
512	0.000477	0.000716	0.001047	0.001425	0.001550	0.001724	0.001824
	0.9970	0.9721	0.9713	0.9649	0.9488	0.9346	0.9036
1024	0.000239	0.000365	0.000534	0.000730	0.000803	0.000902	0.000975
	0.9984	0.9938	0.9928	0.9931	0.9897	0.9743	0.9339

Example 6.2: Consider the following singularly perturbed problem:

$$\varepsilon u_{xx} + a(x)u_x + 2u(x-1,t) - 5u(x,t) - u_t = f(x,t), \quad (x,t) \in (0,2) \times (0,2],$$
$$u(x,0) = 0, \quad x \in [0,2],$$
$$u(x,t) = 0, \quad (x,t) \in [-1,0] \times [0,2],$$
$$u(2,t) = 0, \quad t \in [0,2]$$

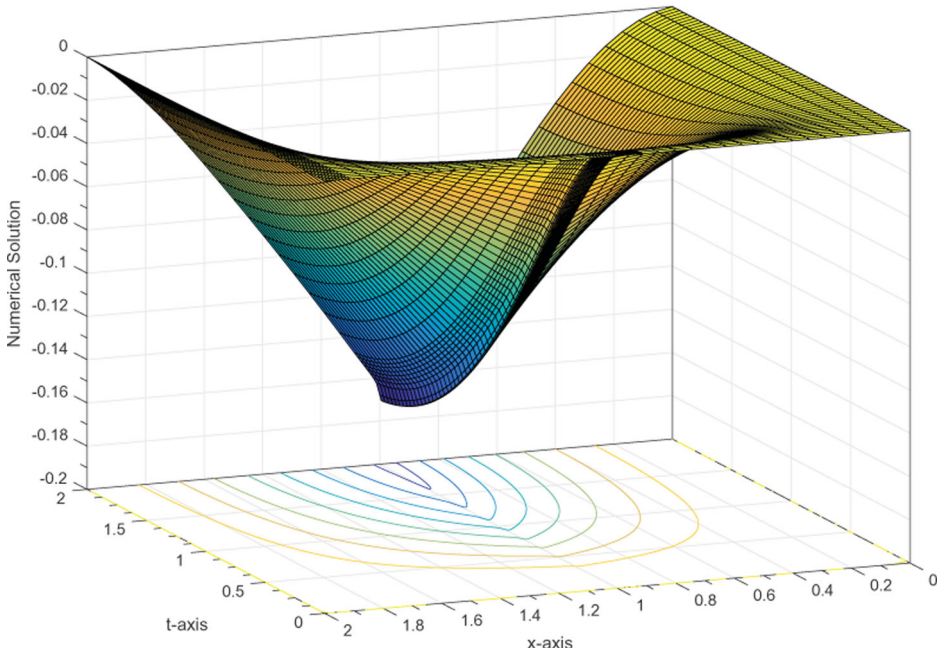


Figure 3. Numerical solution of Example 6.2 for $\varepsilon = 2^{-6}$ when $M = N = 64$.

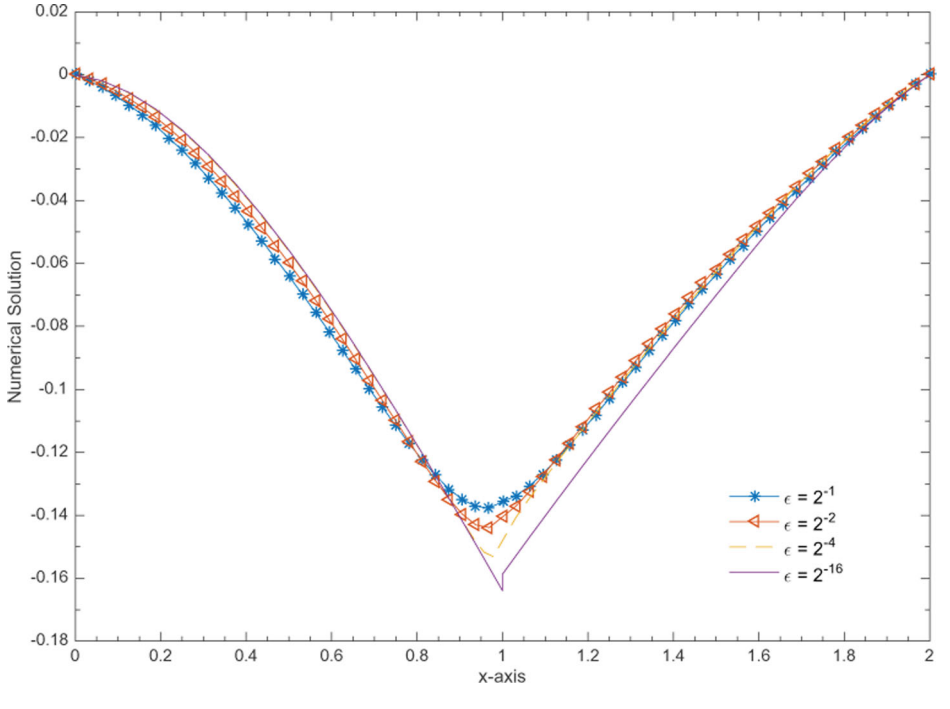


Figure 4. Numerical solutions of Example 6.2 at $t = 2$ for different values of ϵ when $N = 64$.

Table 2. Maximum absolute error and order of convergence for Example 6.1 for different values of ϵ, M, and N when $M = 2N$.

N	$\varepsilon = 2^{-0}$	2^{-1}	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}
32	0.004733	0.007924	0.012284	0.016524	0.018423	0.019067	0.019182
	0.9713	0.9124	0.9063	0.8804	0.8825	0.8848	0.8796
64	0.002414	0.004210	0.006554	0.008976	0.009993	0.010326	0.010426
	0.9857	0.9481	0.9329	0.9622	0.8555	0.8931	0.8890
128	0.001219	0.002182	0.003433	0.004607	0.005523	0.005560	0.005630
	0.9917	0.9660	0.9508	0.9246	0.9811	0.9129	0.9115
256	0.000613	0.001117	0.001776	0.002427	0.002798	0.002953	0.002993
	0.9976	0.9731	0.9584	0.9515	0.9444	0.9364	0.9336
12	0.000307	0.000569	0.000914	0.001255	0.001454	0.001543	0.001567
	0.9953	0.9874	0.9781	0.9783	0.9705	0.9567	0.9479
1024	0.000154	0.000287	0.000464	0.000637	0.000742	0.000795	0.0008123
	0.9967	0.9905	0.9893	0.9884	0.9815	0.9781	0.9721

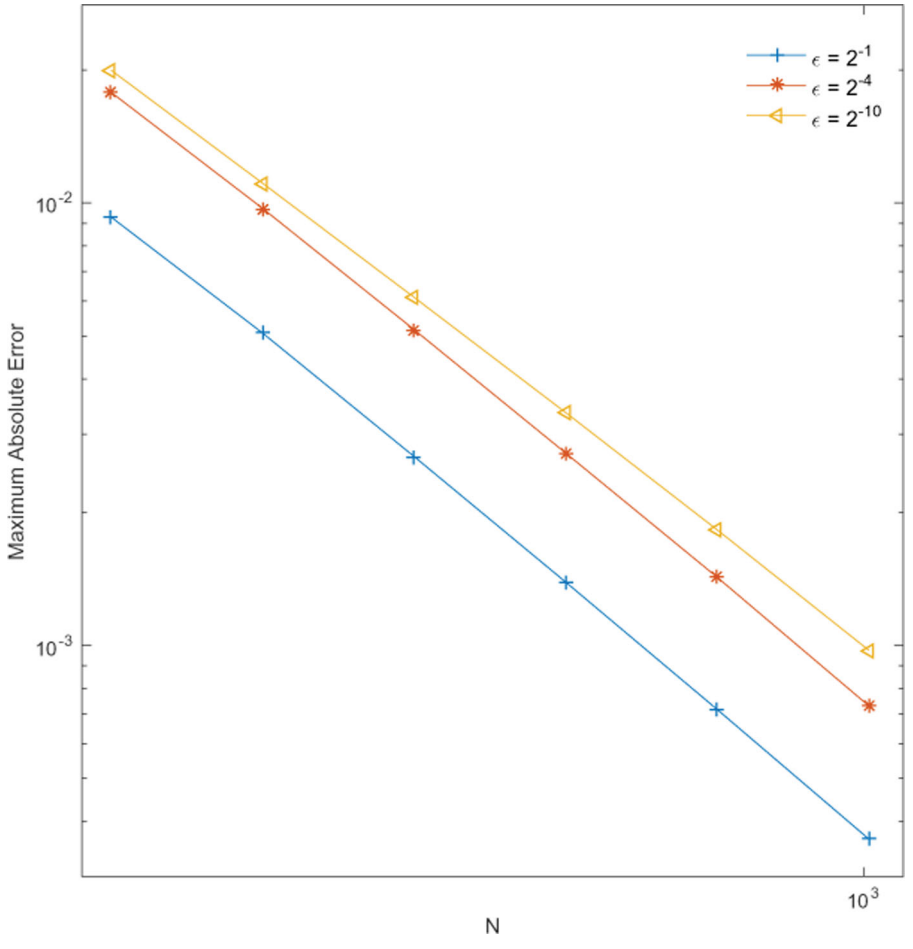


Figure 5. Error plot for Example 6.1.

Table 3. Maximum absolute error and order of convergence for Example 6.2 for different values of ϵ, M, and N when $M = N$.

N	$\varepsilon = 2^{-0}$	2^{-1}	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}
32	0.003389	0.003688	0.003854	0.003972	0.005890	0.005892	0.005894
	0.8167	0.9260	0.9037	0.9017	0.8897	0.8825	0.8749
64	0.001924	0.001941	0.002060	0.002126	0.003179	0.003196	0.003214
	0.9527	0.9184	0.9038	0.8953	0.8853	0.8805	0.8863
128	0.000994	0.001027	0.001101	0.001143	0.001721	0.001736	0.001746
	0.9487	0.9274	0.9197	0.9130	0.8957	0.8881	0.8841
256	0.000515	0.000540	0.000582	0.000607	0.000925	0.000938	0.000946
	0.9586	0.9475	0.9275	0.9236	0.9049	0.9048	0.9027
512	0.000265	0.000280	0.000306	0.000320	0.000494	0.000501	0.000506
	0.9730	0.9694	0.9445	0.9469	0.9250	0.9297	0.9277
1024	0.000135	0.000143	0.000159	0.000166	0.000260	0.000263	0.000266
	0.9843	0.9795	0.9674	0.9672	0.9512	0.9459	0.9417

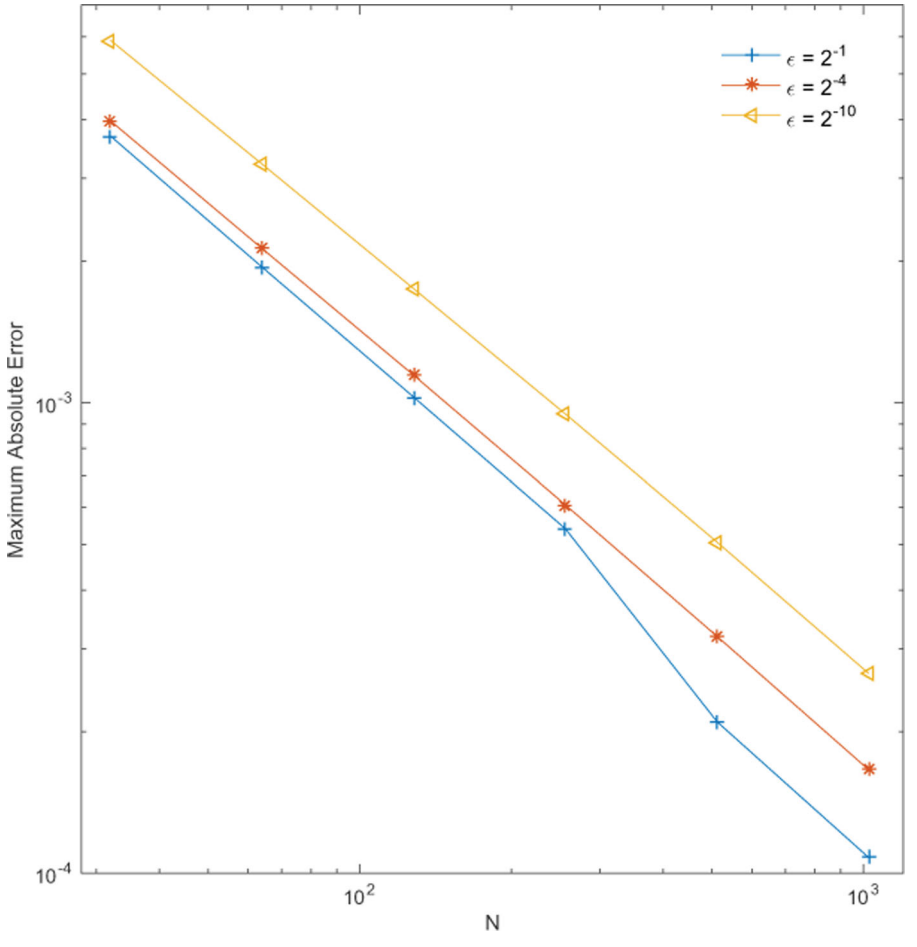


Figure 6. Error plot for Example 6.2.

Table 4. Maximum absolute error and order of convergence for Example 6.2 for different values of ϵ, M, and N when $M = 2N$.

N	$\varepsilon = 2^{-0}$	2^{-1}	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}
32	0.002107	0.002316	0.002723	0.003246	0.003356	0.003378	0.003410
	0.9429	0.9271	0.9191	0.9112	0.9012	0.8859	0.8799
64	0.001096	0.001218	0.001440	0.001726	0.001797	0.001828	0.001853
	0.9713	0.9488	0.9353	0.9330	0.9196	0.9038	0.9029
128	0.000559	0.000631	0.000753	0.000904	0.000950	0.000977	0.000991
	0.9871	0.9661	0.9491	0.9375	0.9289	0.9182	0.9110
256	0.000282	0.000323	0.000390	0.000472	0.000499	0.000517	0.000527
	1.0000	0.9778	0.9635	0.9460	0.9405	0.9319	0.9175
512	0.000141	0.000164	0.000200	0.000245	0.000260	0.000271	0.000279
	1.0102	1.0000	0.9856	0.9708	0.9563	0.9529	0.9442
1024	0.000070	0.000082	0.000101	0.000125	0.000134	0.000140	0.000145
	1.0104	1.0010	0.9921	0.9914	0.9784	0.9749	0.9668

where

$$a(x) = \begin{cases} -(4 + x^2), & x \in [0, 1], \\ (8 - x^2), & x \in (1, 2] \end{cases} \quad \text{and}$$

$$f(x, t) = \begin{cases} 4xt^2e^{-t}, & (x, t) \in [0, 1] \times [0, 2], \\ 4(2 - x)t^2e^{-t}, & (x, t) \in (1, 2] \times [0, 2]. \end{cases}$$

7. Concluding remark

A class of singularly perturbed parabolic partial differential equations with discontinuous coefficients and source terms is solved numerically. The solution of the problem takes into account not just the present state of the physical system but also its history. The simultaneous presence of discontinuous data and delay makes the problem stiff. In the limiting case, the solution of the problem exhibits multi-scale character [21,29]. There are narrow regions where solution derivatives are exponential, and exhibit turning point behaviour across discontinuities besides weak boundary layers.

An implicit numerical scheme based on the upwind finite difference method is presented on a specially generated mesh. The mesh has been chosen such that most of the mesh points remain in the regions with rapid transitions. The proposed numerical method has been analysed for consistency, stability and convergence. Extensive theoretical analysis is performed to obtain consistency and error estimates. It is found that the method proposed is unconditionally stable and the convergence obtained is parameter uniform. Numerical examples have been presented to demonstrate the effectiveness of the technique. Convergence obtained in practical satisfies theoretical predictions. The method presented is easy to implement and, with a little modification, can easily be extended to even more general situations like problems in higher dimensions and the nonlinear evolution equations.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] A.R. Ansari, S.A. Bakr, and G.I. Shishkin, *A parameter robust finite difference method for singularly perturbed delay parabolic partial differential equations*, J. Comput. Appl. Math. 205 (2007), pp. 552–566.

[2] C. Clavero, J.L. Gracia, G.I. Shishkin, and L.P. Shishkina, *Grid approximation of a singularly perturbed parabolic equation with degenerating convective term and discontinuous right hand side*, Int. J. Numer. Anal. Model. 10 (2013), pp. 795–814.

[3] C. Clavero, J.L. Gracia, G.I. Shishkin, and L.P. Shishkina, *Schemes convergent ϵ-uniformly for parabolic singularly perturbed problems with a degenerating convective term and a discontinuous source*, Math. Model. Anal. 20 (2015), pp. 641–657.

[4] C. Clavero, J.L. Gracia, G.I. Shishkin, and L.P. Shishkina, *An efficient numerical scheme for 1d parabolic singularly perturbed problems with an interior and boundary layers*, J. Comput. Appl. Math.318 (2017), pp. 634–645.

[5] M. Chandru, T. Prabha, and V. Shanthi, *A parameter robust higher order numerical method for singularly perturbed two parameter problems with non-smooth data*, J. Comput. Appl. Math. 309 (2017), pp. 11–27.

[6] R.K. Dunne, E. O'Riordan, and G.I. Shishkin, *A fitted mesh method for a class of singularly perturbed parabolic problems with a boundary turning point*, Comput. Methods Appl. Math. 3 (2003), pp. 361–372.

[7] C. de Falco and E. O'Riordan, *Interior layers in a reaction diffusion equation with a discontinuous diffusion coefficient*, Int. J. Numer. Anal. Model. 7 (2010), pp. 444–461.

[8] P.A. Farrell, A.F. Hegarty, J.J.H. Miller, E. O'Riordan, and G.I. Shishkin, Robust Computational Techniques for Boundary Layers, Chapman and Hall/CRC, Boca Raton, 2000.

[9] P.A. Farrell, J.J.H. Miller, E. O'Riordan, Singularly perturbed differential equations with discontinuous source terms, in Proc. of Analytical and Numerical Methods for Convection Dominated and Singularly Perturbed Problems, J.J.H. Miller, G.I. Shishkin, and L. Vulkov, eds., Nova Science Publishers, Inc, New york, 2000, pp. 23–32.

[10] J.L. Gracia and E. O'Riordan, *A singularly perturbed convection diffusion problem with a moving interior layer*, Int. J. Numer. Anal. Model. 9 (2012), pp. 823–843.

[11] J.L. Gracia and E. O'Riordan, *A singularly perturbed parabolic problem with a layer in the initial condition*, Appl. Math. Comput. 219 (2012), pp. 498–510.

[12] P.W. Hemker and G.I. Shishkin, *Discrete approximation of singularly perturbed parabolic PDEs with a discontinuous initial condition*, Comput. Fluid Dyn. J. 2 (1994), pp. 375–392.

[13] A. Kaushik, *Error estimates for a class of partial functional differential equation with small dissipation*, Appl. Math. Comput. 226 (2014), pp. 250–256.

[14] A. Kaushik and M. Sharma, *A robust numerical approach for singularly perturbed time delayed parabolic partial differential equations*, Comput. Math. Model. 23 (2012), pp. 96–106.

[15] A. Kaushik, K.K. Sharma, and M. Sharma, *A parameter uniform difference scheme for parabolic partial differential equation with a retarded argument*, Appl. Math. Model. 34 (2010), pp. 4232–4242.

[16] N. Kopteva and M. Stynes, *Stabilised approximation of interior layer solutions of a singularly perturbed semilinear reaction diffusion problem*, Numer. Math. 119 (2011), pp. 787–810.

[17] T. Linß, Layer Adapted Meshes for Reaction Convection Diffusion Problems, Springer, Berlin, 2010.

[18] J.J.H. Miller, E. O'Riordan, and G.I. Shishkin, Fitted Numerical Methods for Singular Perturbation Problems, World Scientific, Singapore, 2012.

[19] K. Mukherjee and S. Natesan, *ϵ-uniform error estimate of hybrid numerical scheme for singularly perturbed parabolic problems with interior layers*, Numer. Algorithms 58 (2011), pp. 103–141.

[20] N.N. Nefedov, L. Recke, and K.R. Schneider, *Existence and asymptotic stability of periodic solutions with an interior layer of reaction advection diffusion equations*, J. Math. Anal. Appl. 405 (2013), pp. 90–103.

[21] R.E. O'Malley, Introduction to Singular Perturbations, Academic Press, New York, 1974.

[22] E. O'Riordan and J. Quinn, *A singularly perturbed convection diffusion turning point problem with an interior layer*, Comput. Methods Appl. Math. 12 (2012), pp. 206–220.
[23] E. O'Riordan and J. Quinn, *A linearised singularly perturbed convection diffusion problem with an interior layer*, Appl. Numer. Math. 98 (2015), pp. 1–17.
[24] E. O'Riordan and G.I. Shishkin, *Singularly perturbed parabolic problems with non-smooth data*, J. Comput. Appl. Math. 166 (2004), pp. 233–245.
[25] M. Paramasivam, J.J.H. Miller, and S. Valarmathi, *Parameter uniform convergence for a finite difference method for a singularly perturbed linear reaction diffusion system with discontinuous source terms*, Int. J. Numer. Anal. Model. 11 (2014), pp. 385–399.
[26] J. Quinn, *A numerical method for a nonlinear singularly perturbed interior layer problem using an approximate layer location*, J. Comput. Appl. Math. 290 (2015), pp. 500–515.
[27] S.C.S. Rao and S. Chawla, *Interior layers in coupled system of two singularly perturbed reaction–diffusion equations with discontinuous source term*, Lect. Notes Comput. Sci. 8236 (2013), pp. 445–453.
[28] S.C.S. Rao and S. Chawla, *Numerical solution of singularly perturbed linear parabolic system with discontinuous source term*, Appl. Numer. Math. 127 (2018), pp. 249–265.
[29] H.-G. Roos, M. Stynes, and L. Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations, Springer, Berlin, 2008.
[30] V. Subburayan, *A parameter uniform numerical method for singularly perturbed delay problems with discontinuous convection coefficient*, Arab. J. Math. Sci. 22 (2016), pp. 191–206.
[31] V. Subburayan and R. Mahendran, *An ε-uniform numerical method for third order singularly perturbed delay differential equations with discontinuous convection coefficient and source term*, Appl. Math. Comput. 331 (2018), pp. 404–415.

Sensors & Transducers, Vol. 244, Issue 5, September 2020, pp. 37-43

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2020
http://www.sensorsportal.com

Bot Detection in Social Networks Based on Multilayered Deep Learning Approach

[1] Sandeep Singh Sengar, [2] Sanjay Kumar, [2] Pradyot Raina and [2] Mukul Mahaliyan

[1] Department of Computer Science and Engineering, SRM University-AP, Andhra Pradesh, India-522502

[2] Department of Computer Science and Engineering, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, New Delhi, India-110042

E-mail: sandeep.iitdhanbad@gmail.com, sanjay.kumar@dtu.ac.in, pradyotrainaout@gmail.com, mukulmahaliyan@gmail.com

Received: 18 July 2020 /Accepted: 28 August 2020 /Published: 30 September 2020

Abstract: With the swift rise of social networking sites, they have now come to hold tremendous influence in the daily lives of millions around the globe. The value of one's social media profile and its reach has soared highly. This has invited the use of fake accounts, spammers and bots to spread content favourable to those who control them. Thus, in this project we propose using a machine learning approach to identify bots and distinguish them from genuine users. This is achieved by compiling activity and profile information of users on Twitter and subsequently using natural language processing and supervised machine learning to achieve the objective classification. Finally, we compare and analyse the efficiency and accuracy of different learning models in order to ascertain the best performing bot detection system.

Keywords: Bot detection, Machine learning, Natural Language Processing, Social network, Text classification.

1. Introduction

Today, social networking sites have acquired an integral position in the daily lives of almost every person with access to the internet around the globe. The amount of time the average user spends on social media apps has also been increasing at an almost alarming rate for some time. Nowadays, users don't just use social networks like Facebook, Twitter, Instagram, etc. for interacting within their social circles but also for getting news updates, multimedia entertainment, political discussions, business and even shopping. All this points to the fact that social media has a massive hold on a significant part of the world's population and can thus be used, or misused, for influencing them. With people increasingly resorting to the social media profiles to judge and make views of the people around them, it has become increasingly important to have an impressive social media following. This is particularly a cause of concern for celebrities, politicians and others with similar public careers. Moreover, social media has the tendency to broadcast content to millions of people in a matter of minutes and thus all prominent organizations, from political parties to corporations, have started using it to push their own narratives and sway public opinion one way or another. And so, the use of fake accounts and bots in an organized manner on a massive scale has recently become a short cut to becoming 'viral' and inflating one's social media presence. 'Bots' are simply computer programs that automatically produce or repost content and interact with humans on social

networks. When used on a large enough scale, such bots have had significant impact on the real world – from spreading fake news during elections, influencing stock prices, swaying opinion about a company or a product, hacking and cybercrimes, data stealing and also promoting and distorting the popularity of individuals and groups. Thus, with our project we aim to identify and distinguish bots from genuine humans in order to curb the menace caused by them. To achieve this, we have proposed a layered machine learning approach. Starting with a labelled dataset, we perform feature extraction based on established tests and correlation analysis. Then, text attributes are processed using feature engineering followed by initial classification that produces a vector of predictions for each text attribute. We further provide a comparative analysis of the performance of various well-known classification algorithms based on their prediction accuracy on the testing dataset. This evaluation also includes a comparison of two popular feature engineering techniques for text attributes i.e. Bag-of-Words and n-gram model.

1.1. Dataset

Our project has been implemented on a Twitter dataset which contains information about more than 5000 different Twitter users and nearly 200,000 tweets. The user information is across attributes including their *handle, name, location, followers count, profile description or bio, favourites count, tweets count, sample tweets,* etc., which we compiled together to build our model. Out of these 20 attributes, 16 attributes are either numerical or categorical while the remaining 4 i.e. *description, status, name* and *screen_name,* have text values. Also, the dataset includes labels for each of its instances, classifying a user as a bot (1) or a genuine user (0) which enabled us to take a supervised learning approach and train classifier algorithms over it.

The approach proposed in this paper generates predictions based entirely on a user's public data and information which can be easily extracted from Twitter itself in real-time.

Table 1. Comparative analysis of the related work.

Author, Year	Dataset	Approach	Description	Advantages	Disadvantages
Wei F., *et al*., 2020	Cresci, 2017	Model	Classification using BiLSTM with word embeddings	Comprehensive tweet level analysis using RNN	Does not take into account user level data
Luo L., *et al*., 2020	PAN Bots and Gender Profiling	Model	Deepbot consisting of a classifier and a Web interface	Uses GloVE embedding + BiLSTM for better tweet analysis	Does not take into account user level data
Kudugunta S., *et al*., 2018	Cresci, 2017	Layered Model	Two-layered model for classification	Combines tweet level and tweet text analysis for more reliable result	Does not account for both textual analysis and account level analysis
Zhao C., *et al*., 2020	Chen, *et al*., 6 million spam tweets	Algorithm	Uses a layered ensemble framework	Better performance than normal LSTM models on tweet level data	Only takes into account tweets, not user data

2. Literature Review

The literature of twitter bot detection is vast, but the problem is still a very challenging task with bots becoming smarter with time. It's relatively simple to detect bots which show unusual behaviour such as retweeting very frequently. But smarter bots curb to show unusual behaviour and act more humanly. It is harder for bots to show similar tweeting patterns as humans thus deep learning models are quite successful in detecting bot by textual analysis.

The existing state of research in this field consists of approaches such as "Twitter Bot Detection Using Bidirectional Long Short-term Memory Neural Networks and Word Embeddings. 2019" and "Deepbot: A Deep Neural Network based approach for Detecting Twitter Bots" employ latest deep learning techniques like Bi-LSTM based approach for textual analyses for bot detection. A similar approach "Kudugunta S., Ferrara E.: Deep Neural Networks for Bot Detection" models the user classification task differently at user and tweet level data and comparisons based on performance of each level are made.

Our work combines the two levels and attempts to find unusual behaviour in engagement of the user and also in tweets of the user. Bi-LSTMs are slightly better than LSTMs in the sense that each state has memory of not just previously occurring series but also the memory of future occurring series. LSTMs perform good on time series data where sequence of the sentence is important. CNN has performed equally good or even better than LSTMs in text classification tasks "Yin W., Kann K., Yu M., Schütze H.: Comparative Study of CNN and RNN for Natural Language Processing. arXiv". when feature detection is extremely important, and certain sequence to sequence learning tasks "Gehring J., Auli M., Grangier D., Yarats D., Dauphin Y.N.: Convolutional sequence to sequence learning. 34th Int. Conf. Mach. Learn.".

In the context of this problem, identification of features such as abuses, angry terms, and hate speech

is very useful in understanding the context of the tweet. Thus, we employed CNN for textual analyses of tweets. Our proposed approach achieves performance as good as or better when compared to existing work "Twitter Bot Detection Using Bidirectional Long Short-term Memory Neural Networks and Word Embeddings. 2019", "Deepbot: A Deep Neural Network based approach for Detecting Twitter Bots", "Kudugunta, S., Ferrara, E.: Deep Neural Networks for Bot Detection" and "Zhao, C., Xin, Y., Li, X., Yang, Y., Chen, Y.: A Heterogeneous Ensemble Learning Framework for Spam Detection in Social Networks with Imbalanced Data. Appl. Sci. 10, 936 (2020)".

3. Methodology

Before performing the actual learning on our data, we analyse our dataset in a comprehensive manner in order to better identify strongly correlated attributes, highlight weak correlations and determine which features are most significant in the prediction of the target variable i.e. *bot* status. Due to the variance in attribute types we proceed in steps towards our aim of generating a smaller set of strong features to train our learning algorithm on.

First, we separate the 4 text valued attributes and perform feature engineering on each of them in a subsequent step. Of the remaining non-text attributes, by elementary analysis tools and intuition, we are able to eliminate some like *location, created_at,* etc. as these contain either majority NaN, have a single repeated value, are completely random or strongly correlated with another attribute in the dataset. Thus, on the remaining numerical/categorical attributes and 4 text attributes, we perform feature engineering. For the non-text attributes, we obtain the following correlation matrix, as shown in Fig. 1 that provides an approximate idea about the relation between the attributes and the target variable i.e. *bot.* However, feature selection based purely on correlation is not always reliable. Thus, we apply the chi^2 test to determine the 5 best or strongest features and the resulting dataset with reduced dimensionality includes the attributes *id, followers_count, friends_count, listed_count* and *statuses_count*.

Fig. 1. Correlation matrix for user-level non text attributes.

Having previously separated the 4 text attributes, we proceed with feature engineering for these by observing the correlation between them and the target variable one by one. To achieve this, we employ two feature extraction methods i.e. Bag-of-Words model and the n-gram model and then compare results for the

two. Once we derive the 'features' or 'vectors', which are simply the frequency of words appearing in the text after the appropriate transformations, we train a classifier on each of these 4 vectorized attributes and use it to make predictions for the training samples which are then recorded in a column. The resulting 4 prediction columns, one for each text attribute, are then added to the selected features.

In the final step, we perform the final classification using the extracted feature set and the *bot* labels by training different classification models and then comparing their accuracy scores. This step is performed for two feature sets i.e. one obtained from Bag-of-Words extraction and the other from n-gram extraction technique.

As stated earlier, we follow a supervised learning approach to learn how a bot and human differ in the behaviour on social networks. First, we load the data into a *pandas* dataframe to make it usable in the Python environment. By analysis we find that there are some missing values throughout the dataset so we proceed by filling these values in order to prevent errors. This leads us to the feature selection and extraction step where we separate the text attributes beforehand. On the remaining dataset consisting only of numerical and categorical attributes, we select a few best performing attributes based on the chi^2 test. Now, we turn our focus to extracting useful features from the text attributes. The text attributes to be analysed are *screen_name*, *name*, *description* and *status*. The *status* attribute contains nearly 200k tweets grouped by the user id of the user Now, we employ two different feature engineering techniques i.e. Bag-of-Words and n-gram.

These methods work by analysing how frequently each word occurs and how this relates to the target attribute. Once this 'vectorization' is achieved, we perform term frequency-inverse document frequency transformation in order to normalize the 'vectors' and pre-empt bias towards higher frequency words. The result after this is a sparse matrix of transformed word frequencies which we then convert to pandas dataframes for all the four attributes and then concatenate them together to perform classification for all four together. Also, the feature sets resulting from the Bag-of-Words model and the one resulting from the n-gram model are maintained separately in order to allow comparisons at a later stage.

The flowchart of the methodology followed in the implementation of this research paper is elaborated in Fig. 2.

Once the two step feature extraction process is finished and the final training dataset is generated, we perform the actual learning using several different established machine learning algorithms in order to compare and analyze their performance and determine the best suited model for our problem. The machine learning algorithms thus used and analyzed for performing the classification in this paper are

- K Nearest Neighbours Classifier
- Decision Tree Classifier
- Random Forest Classifier
- Ada Boost Classifier
- Gradient Boosting Classifier
- Gaussian NB Classifier
- Multinomial Naïve Bayes Classifier, and
- Multilayer Perceptron

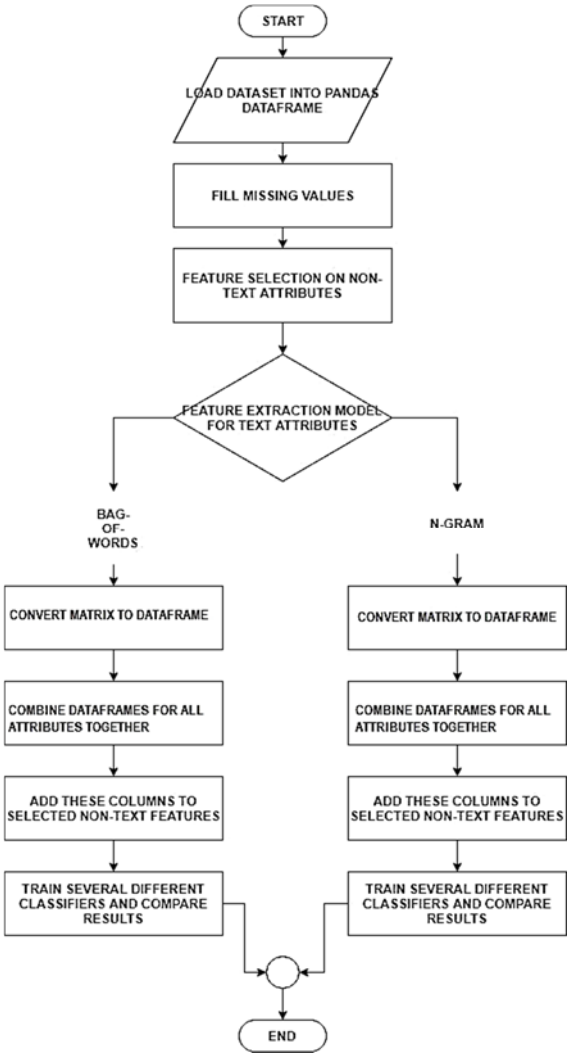


Fig. 2. Flowchart representing the methodology of the proposed method.

4. Findings

The accuracy scores for different classifiers on both the textual feature extraction models are detailed in Table 2 and Table 3 below.

As seen, the multilayer perceptron and multinomial Bayes classifiers show the best results when the bag-of-words model is used to extract features from the textual data.

As is evident from the table of accuracy scores above, the n-gram model of feature extraction offers significantly better accuracy on average over different classifiers with 89.7 % average accuracy than the Bag-of-Words model with average accuracy score of 87.3 %.

Table 4 below details an analysis of the performance of each classifier.

Table 2. Accuracy Scores for bag-of-words model.

S. No.	Algorithm	Accuracy Score
1.	K Neighbors Classifier	57.1019%
2.	Decision Tree Classifier	92.9699%
3.	Random Forest Classifier	86.8006%
4.	Ada Boost Classifier	91.1047%
5.	Gradient Boosting Classifier	91.6069%
6.	Gaussian Naïve Bayes	89.3113%
7.	Multinomial Naïve Bayes	94.7633%
8.	Multilayer Perceptron	94.7633%
	Average	**87.3027%**

Table 3. Accuracy scores for n-gram model (n=2).

S. No.	Algorithm	Accuracy Score
1.	K Neighbors Classifier	79.5552%
2.	Decision Tree Classifier	92.8264%
3.	Random Forest Classifier	92.6112%
4.	Ada Boost Classifier	90.6026%
5.	Gradient Boosting Classifier	90.0287%
6.	Gaussian Naive Bayes	87.7331%
7.	Multinomial Naïve Bayes	89.3831%
8.	Multilayer Perceptron	95.4089%
	Average	**89.7685%**

Table 4. Accuracy scores(%) of each classifier model.

S.No	Classifier	Bag-of-Words accuracy	n-gram accuracy	Average Accuracy
1.	K Neighbors	57.10%	79.55%	**68.32%**
2.	Decision Tree	92.96%	92.82%	**92.89%**
3.	Random Forest	86.80%	92.61%	**89.70%**
4.	Ada Boost	91.10%	90.60%	**90.85%**
5.	Gradient Boosting	91.60%	90.02%	**90.81%**
6.	Gaussian Naïve Bayes	89.31%	87.73%	**88.52%**
7.	Multinomial Naïve Bayes	94.76%	89.38%	**92.07%**
8.	Multilayer Perceptron	94.76%	95.40%	**95.08%**
	Overall Average			***88.53%***

5. New Findings

There are many deep learning algorithms available that could be applied to text classification. LSTMs are highly useful in case of time series data. LSTMs evaluate the relationship between the consecutive time step data. LSTMs as opposed to RNNs keep a memory of previously occurring series (which in current case is the sequence of words) which it uses while feed forwarding from cell to cell. However, LSTM has a few demerits in the context of the problem. LSTMs are not the best model where feature detection is paramount. For example searching for abuses, angry terms and named entities, ConvNet works better. LSTMs work better in tasks where sequence of the sentence is important, which is seen in problems of translation using computers or automated answering to queries. Studies in "Yin W., Kann K., Yu M., Schütze H.: Comparative Study of CNN and RNN for Natural Language Processing. arXiv" show clear comparison. Thus, we select ConvNet as the text classification model.

For using ML or deep learning algorithms on texts the necessary step is to apply appropriate language modelling technique. Language modelling allows representation of textual data in a vectorized form which has a high imperative in machine learning. We used Glove "Pennington J., Socher R., Manning C. D.: Glove: Global Vectors for Word Representation. 1532–1543 (2014)" word embedding for language modelling which is a better technique than previously used language modelling techniques namely Bag-of-words and N-gram. Glove is a context independent open source embedding provided by Stanford University that captures semantic similarity between large vocabularies of words. It is also useful for handling out-of-dictionary words which were absent in the dataset during deployment of the model. ConvNets are usually applied to computer vision problems but also show promising results on NLP problems.

The input to the convolutional neural network model used during tweet level analysis to generate the first level of predictions based on the textual data can be mathematically represented as:

Input dimensions =

=(No of words in text) × (Embedding length)

Fig. 3 represents a dummy input matrix having an embedding length of 5 and text "I am a Bot #kill #lol !".

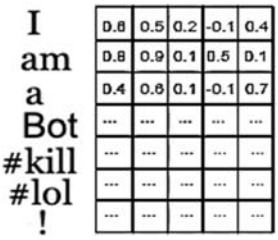


Fig. 3. A dummy input matrix to the CNN text classification model.

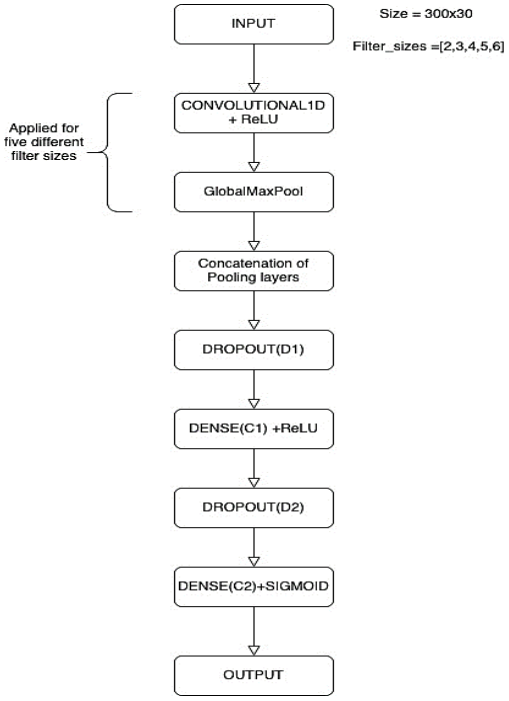


Fig. 4. Architecture of the proposed CNN used for tweet level analysis.

Row 1 corresponds to embedding weights of word number 1, in the input sentence which is 'I', as stored in Glove word embeddings. Similarly, Row 4 corresponds to embedding weights of word number 4 in the input sentence which is 'Bot' in this case and so on.

5.1. Algorithm

Let set U contain U1, U2 … Un Ui represents a user tuple with features uf(1), uf(2) … uf(n-1), uf(n). // total n features and set T contains T1, T2 … Tn, where Ti represent a tweet tuple with features tf(1), tf(2) … tf(m-1), tf(m) and tf(text). // total m+1 features.

Step 1:
Select an appropriate k.
Select a random Tk tweet from T and discard the all Ui in U where Ui.userid == Tj.userid (for all Tj int Tk).

Step 2:
Train CNN text classifier with all tf(text) in Tk. Let C(x) be the function that describes this classifier. Discard Tk from T. Generate a new features tf(CNN prediction) in T calculated as:
Ti.tf(CNN prediction) = C(Ti.tf(text))

Step 3:
Generate new features uf(n+1) … uf(n+m+1) in U which corresponds to tf(1) … tf(m) and tf(CNN prediction).

Step 4:
For all tuples Ti, Ti+1 ... Ti+j in T, where Ti.userid = Un.userid calculate the central tendency using A.M. for each feature tf(j). Assign this to uf(n+j).

Step 5:
Split U in train_set and test_set

Step 6:
Choose an appropriate machine learning classifier which is used to train on train_set. Let M(x) describe this classifier.

Step 7:
M makes predictions for each tuple in test_set and performance of the classifier is evaluated on the basis of disparity in prediction and original value.

5.2. Results

The model was trained and tested on very large data to avoid sampling bias and obtain more reliable results. The dataset was acquired from MIB Datasets (Cresci), The dataset consists of a total of 8,386 accounts and 62,96,505 tweets of the accounts. The new approach gave a staggering improvement of prediction result and gave an average 97.033 % prediction over an average of 88.5 % having used an extremely larger dataset than the latter.

Table 2. Results obtained by combining user data with tweet metadata and CNN predictions using the proposed new approach.

Classifier	Accuracy	Precision	Recall	F-score
K Neighbours	96.10	96.09	96.10	96.09
Support Vector	94.78	94.76	94.78	94.60
Decision Tree	99.48	99.48	99.48	99.48
Random Forest	**99.54**	99.54	99.54	99.54
Ada Boost	99.48	99.48	99.48	99.48
Gradient Boosting	**99.54**	99.54	99.54	99.54
Gaussian NB	90.31	90.76	90.31	89.24
Linear Discriminant	97.42	97.71	97.42	97.47
Multilayer Perceptron	96.67	96.87	96.67	96.73

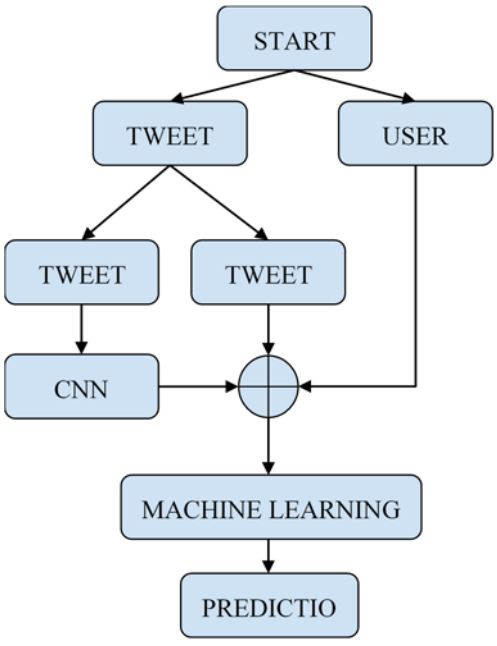


Fig. 5. An abstraction of the newly proposed model.

6. Conclusions

This research paper lays emphasis on the identification, detection and eradication of bots that populate social networks today. We focus on Twitter to determine which user is a bot and which is a human based on their activity and profile information. By taking a machine learning approach for such

distinction, we are not only able to detect bots with sufficient accuracy but this also enables us to define the behaviour and activity patterns exhibited by bots. Apart from the mechanism to detect bots, we also offer a comparative analysis of the performances of various established machine learning techniques and algorithms in our project. We evaluate the accuracy of several classification algorithms when applied after two different feature engineering techniques. The results show that neural networks based Multilayer Perceptron algorithm gives the most accurate predictions. We find that our proposed technique results in an average accuracy of 88.5 % across several established tools.

Later research observed that using convolutional neural networks for text level prediction of the stacked ensemble the results flung very high and showed an average of 97.03 %.

References

[1]. Oentaryo Richard J., Murdopo Arinto, Prasetyo Philips K., Lim Ee-Peng, On Profiling Bots in Social media, in *Proceedings of the International Conference on Social Informatics*, 2016.

[2]. Sandeep Singh Sengar, Sanjay Kumar, Pradyot Raina, Mukul Mahaliyan, Bot Detection in Social Networks Based on Machine Learning Techniques, User Information and Activities, in *Proceedings of the 2nd International Conference on Advances in Signal Processing and Artificial Intelligence*, Berlin, Germany, 2020, pp. 121-124.

[3]. Fazil Mohd, Abulaish Muhammad, A Hybrid Approach for Detecting Automated Spammers in Twitter, *IEEE Transactions on Information Forensics and Security*, Vol. 13, Issue 11, 2018, pp. 1-4.

[4]. Aswani Reema, Kar Arpan, Ilavarasan Vigneswara, Detection of Spammers in Twitter marketing: A Hybrid Approach Using Social Media Analytics and Bio Inspired Computing, *Information Systems Frontiers*, Vol. 20, Issue 2, 2017, pp. 1-16.

[5]. Sengar S. S., Mukhopadhyay S., Moving object detection based on frame difference and W4, *Signal, Image and Video Processing*, Vol. 11, Issue 7, 2017, pp. 1357-1364.

[6]. Efthimion Phillip George, Scott Payne, Nicholas Proferes, Supervised machine learning bot detection techniques to identify social twitter bots, *SMU Data Science Review*, Vol. 1, Issue 2, 2018.

[7]. Sengar S. S., Mukhopadhyay S., Moving object area detection using normalized self adaptive optical flow, *Optik*, Vol. 127, Issue 16, 2016, pp. 6258-6267.

[8]. Wang A.H., Machine Learning for the Detection of Spam in Twitter Networks, in: Obaidat M. S., Tsihrintzis G. A., Filipe J. (Eds.), e-Business and Telecommunications. ICETE 2010. Communications in Computer and Information Science, *Springer*, Berlin, Heidelberg. Vol. 222, 2012, pp. 319-333.

[9]. A. Mostrous, M. Bridge, K. Gibbons, Russia used Twitter bots and trolls 'to disrupt' Brexit Vote, 2017. [Online]. Available: https://www. thetimes.co.uk/ edition/news/russia-used-web-posts-to-disrupt-brexitvote-h9nv5zg6c

[10]. A. Bessi, E. Ferrara, Social Bots Distort the 2016 U.S. Presidential Election Online Discussion, Nov. 2016. [Online]. Available: http://firstmonday.org/ojs/ index.php/fm/article/view/7090

[11]. Sengar S. S., Mukhopadhyay S., Motion detection using block based bi-directional optical flow method, *Journal of Visual Communication and Image Representation*, Vol. 49, 2017, pp. 89-103.

[12]. S. Kramer, Identifying Viral Bots and Cyborgs in Social Media, 2017. [Online]. Available: https://www.oreilly.com/ideas/identifying-viral-bots-and-cyborgs-in-social-media

[13]. Sengar S. S., Mukhopadhyay S, Moving object detection using statistical background subtraction in wavelet compressed domain, *Multimedia Tools and Applications*, 2019.

[14]. S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, M. Tesconi, The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race, in *Proceedings of the 26th International Conference on World Wide Web Companion (WWW)*, Geneva, Switzerland, 2017, pp. 963-972.

[15]. Sengar S. S., Mukhopadhyay S., Motion segmentation-based surveillance video compression using adaptive particle swarm optimization, *Neural Computing and Applications*, 2019, pp. 1-15.

[16]. P. Suárez-Serrato, M. E. Roberts, C. Davis, F. Menczer, On the influence of social bots in online protests, in *Proceedings of the International Conference on Social Informatics (SocInfo'2016)*, 2016, pp. 269-278.

[17]. Sengar S. S., Motion segmentation based on structure-texture decomposition and improved three frame differencing, in *Proceedings of the 15th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI'19)*, Hersonissos, Crete, Greece, 2019, pp. 609-622.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2020 (http://www.sensorsportal.com).

Journal of Intelligent Information Systems
https://doi.org/10.1007/s10844-020-00625-6

Community detection in complex networks using network embedding and gravitational search algorithm

Sanjay Kumar[1,2] · **B S Panda**[2] · **Deepanshu Aggarwal**[1]

Received: 2 June 2020 / Revised: 17 August 2020 / Accepted: 11 October 2020 /
Published online: 05 November 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

The structural and functional characteristic features of nodes can be analyzed by visualizing community structure in complex networks. Community detection helps us to detect nodes having similar behavior in a system and organize the network into a network of closely connected groups or modules. Network embedding technique represents the nodes of the input graph into vector space and preserves their inherent and topological features and can contribute significantly to various applications in network analysis. In this paper, we propose a novel community detection method using network embedding technique. Firstly, nodes of the graph are embedded in feature space of d dimensions, and then low-rank approximation is applied to avoid the results from being affected by noise or outliers. Further, k-means clustering is employed to find the centroids of the clusters in the network and followed by a gravitational search algorithm to improve the results of centroids of clusters. Finally, we compute the effectiveness of detected communities using different performance measures. Our method serves as a universal framework towards applying and bench-marking various embedding techniques in graphs for performing community detection. We perform the test using various evaluation criteria on several real-life and synthetic networks and the obtained result reveals the utility of the proposed algorithm.

Keywords Community detection · Complex networks · GSA (gravitational search algorithm) · k-means clustering · Network embedding · Social networks

✉ Sanjay Kumar
sanjay.kumar@dtu.ac.in

B S Panda
bspanda@maths.iitd.ac.in

Deepanshu Aggarwal
deepanshu4929@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India

[2] Computer Science and Application Group, Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

1 Introduction

Most real-life networks, like social networks, biological networks, transportation networks, power networks, collaboration networks are complex networks composed of a large number of active entities (Newman 2001, 2012; Freeman 2004; Saleh et al. 2019). The study of complex networks helps us identify natural and social system dynamics. It has led to an interdisciplinary association between researchers from different fields like physics, mathematics, biology, sociology, computer-science to collaborate, analyze, and reveal interesting features (Mitchell 2006; Aggarwal 2011; Aggarwal and Subbian 2014). A complex network can be represented as a graph $G = (V, E)$, where V denotes nodes or entities in the system, and E corresponds to the relationship between the vertices. For example, in online social networks, nodes are the person, and edges are the virtual acquaintance like friendship, follow-followee. In the case of the airport network in the world, cities are nodes, and an edge between two cities means there is a direct flight between them.

The presence of community structure is one of the inherent features and quite common in real-life networks (Fortunato 2010; Girvan and Newman 2002). In general, a community is defined as a group of densely interconnected nodes and sparsely connected to the rest of the network (Chouchani and Abed 2020). Finding community structure in complex networks reveals many exciting features and association between entities. Community detection is a type of graph partitioning problem with proper optimization require to find the solution and lies under the NP-hard category (Fortunato and Hric 2016). The idea of community detection relies on the similarity between entities having similar properties or interests in the network and associating them into the same cluster or community. Practically this results in similar nodes forming small groups with high connectivity, connected by sparse edges with the nodes of other groups. Communities in social networks correspond to a set of people sharing similar hobbies, interests, backgrounds (Ramezani et al. 2018). In a protein-protein interaction network, a community may refer to modules with similar biological functions (Mahmoud et al. 2013). A community in collaboration networks represents a group of researchers to plan their work in associations to accomplish common goals (Battiston et al. 2016). Hence, community detection helps to comprehend and uncover the functional behavior of such complex systems. Finding communities in complex networks have a wide range of applications. Few examples of these applications include influence maximization and viral marketing (Kumar and Panda 2020; Huang et al. 2019), link prediction in complex networks (Yadav and Rai 2020), identifying opinion leaders (Bamakan et al. 2109), suggesting products or friends in recommendation system (Sahebi and Cohen 2011), and information propagation in networks (Danon et al. 2008).

Community detection is one of the widely researched topics in the field of network science. In the recent years, numerous community detection algorithms have been proposed to reveal community structures. Most of them are based on graph-theoretic approach (Clauset et al. 2004; Blondel et al. 2008) like modularity optimization (Newman 2006a), label propagation (Cordasco and Gargano 2010) and nature-inspired algorithms (Jaradat and Hamad 2018; Messaoudi and Kamel 2019; Honghao et al. 2013; Pattanayak et al. 2019) with certain limitations in each method, and scaling issues for applying in large networks. Most of these conventional techniques work well in the system with a clear community structure, but their performance may deteriorates in a hazy structure network.

A general graph representation like adjacency matrix and adjacency list can be used to represent a complex network. But, for a vast network having millions of nodes, the conventional representation poses several challenges like high computational complexity, low

parallelizability, and inapplicability of machine learning methods. Recently, many graph embedding or network embedding methods are proposed to represent networks in low-level vector representation while preserving the topology and various features of nodes. We exploit such preservation of the topological features of nodes using the network embedding techniques and the notion of clustering in the feature space to uncover community structures. In this paper, we propose a community detection algorithm, which is a combination of network embedding and evolutionary-based methods. We use three state-of-art recently proposed network embedding methods, namely HOP, node2vec, and SDNE, in our algorithm. We obtain the low-dimensional vector representation of each node of the network by above said embedding methods then apply low-rank approximation to avoid the results from being affected by noise or outliers. Then we perform k-means clustering to obtain the centroids of clusters. The results of k-means may trap into local optima. The famous population optimization-based gravitational search algorithm is further employed to improve the results globally to detect community structures. Usually, swarm-based algorithms may suffer from the issue of random selection of the initial population and may fall into the problem of local optima. The output of k-means clustering as one of the initial population to GSA can serve as a valuable input, which is crucial for the GSA algorithm. We execute the proposed algorithm on many real-life and synthetic networks and estimate the effectiveness of identified communities using different performance measures like normalized mutual information (NMI), modularity, precision, and recall and obtained commendable results.

The main contributions of the proposed algorithm are:

- We devise a generic framework to apply various network embedding techniques on graph datasets to uncover community structure in complex networks.
- We introduce a community detection algorithm, which uses the notion of geometric space-based and evolutionary-based methods.
- The outcome of the k-means clustering as one of the input enhances the quality of the initial populations of GSA ensures a globally optimal solution.
- Experimental results obtained on various real-life and synthetic datasets using the different evaluation criteria reveal the effectiveness of the proposed algorithm.

The rest of the paper is organized as follows: Section 2 presents the the various community detection algorithms. Section 3 discusses the preliminaries of the work, including network embedding techniques, clustering methods, and gravitational search algorithm. The detailed descriptions of the proposed algorithm are described in Section 4. Section 5 lists the various real-life and synthetic datasets used for the experiments and the different evaluation criteria to judge the performance. Section 6 reports the results obtained and their analysis. Finally, we conclude the paper in Section 7.

2 Related work

Revealing the community structure manifested by real-life networks is a significant measure for an insight into complex systems, which go beyond the local structure of the members in the network. Due to the importance of community structure in the network analysis, numerous community detection algorithms are proposed by researchers of different disciplines. These methods are majorly partitioned into following three classes: graph theoretic based algorithms (Girvan and Newman 2002; Clauset et al. 2004; Blondel et al. 2008),evolutionary algorithms based techniques (Messaoudi and Kamel 2019; Hanghao et al. 2013; Pattanayak

et al. 2019), and geometric space-based algorithms (Mahmood and Small 2015; Eustace et al. 2014; Ding et al. 2018). Graph-theoretic based algorithms exploit the various network topological features and construct like finding clique, clustering, label propagation, random walk, node centrality based methods, and others. Further such methods can be further subdivided into three categories -node-centric, and network-centric methods. In node-centric methods, nodes should satisfy some specific properties if they belong to a community. Some of these properties are - complete mutuality, reachability of members, node degrees, similarity between nodes. Usually, node centric method finds an initial subset of nodes as the seed nodes in the network and communities are extended from the seed set using techniques like label propagation (Gui et al. 2018; Ding et al. 2018). The node-centric Clique Percolation Method (CPM) rely on pre-defined substructures like finding cliques and k-cores within the network, which extracts all the maximal cliques in the network, to determine communities (Palla et al. 2005). CPM uses k-cliques to determine communities where a k-clique represents a complete subgraph with k nodes. Two k-cliques are said to be adjacent if they have an overlap of $k-1$ nodes. Groups of adjacent k-cliques can form a network community. However, most node-centric methods often suffer from location and seed bias. To overcome such issue Ma et al. (2014) proposed a seed insensitive for local community detection. They measure the similarity among nodes by examining their neighborhoods and report a local community by maximizing internal similarity and minimizing external similarity simultaneously. Rosvall and Bergstrom (2008) proposed an information-theoretic approach known as infomap, a flow-based method that detects communities based on the map equation. Guo et al. (2020) proposed a local community detection algorithm using internal force. They obtain core members of communities as seeds in the network using local degree and Jaccard coefficient, and the fitness function uses the node with the highest degree among seeds every time.

Network-centric techniques rely on techniques like multidimensional scaling, spectral clustering, and modularity maximization (Newman 2006a; Zhou and Amini 2019). Newman (2006a) proposed a fast greedy approach for finding communities in the network which is based on the optimization of the quality function known as modularity. The objective of the partitioning of a network into communities is to maximize the intra-community interactions and minimize the inter-community interactions. In network-centric community detection, the connection of nodes are considered globally then partition the network into disjoint sets based on node similarity, latent space model, and modularity maximization. Based on modularity maximization, one of the popularly used methods for community detection is the Girvan and Newman (GN) Algorithm (Girvan and Newman 2002). It is a divisive hierarchical algorithm that removes links between vertices iteratively based on the value of betweenness of the links. This link removal process terminates upon reaching the maximum modularity of resulting partitions. The modularity formula of Newman and Girvan is a well-known quality function that estimates the goodness of a partition of the graph for community detection. However, the GN method experiences the issue of scalability and high time complexity. Recently, Arasteh and Alizadeh (2019) proposed a faster community detection technique based on the idea of edge degree betweenness centrality, and more than one edge can be removed to find the modular structure in the network using the idea of edge betweenness. The proposed method enhances the performance of the GN method.

In recent years, researchers explored various evolutionary methods to uncover community structure in networks and produced outstanding results. Honghao et al. (2013) introduced the Ant Colony Optimization (ACO) technique inspired by real ant colony formation heuristics for detecting community. Using a novel heuristic information and the

locus-based adjacency presentation of the network, they adopted the max-min ant system (MMAS) framework to identify network community structure. Jaradat and Hamad (2018) proposed the firefly algorithm, where they randomly initialize the fireflies and move them using intensity equations and node connectivity for detecting communities. Authors in (Pattanayak et al. 2019) proposed a fire propagation-based community detection algorithm by relating the seed nodes in the community detection as the fire starting to form a source. They confined the search space by a two-radius neighborhood graph. Generally, evolutionary-based algorithms either set parameters in advance or uses a random method to produce the initial population leading to large search space and high computational complexity. Messaoudi and Kamel (2019) introduced multi-objective Bat Algorithm, which uses the Mean Shift algorithm to generate the initial population to control randomness in initial population choice, to offer better solutions. However, high complexity cost and non-scalability is the main challenge for these evolutionary algorithms.

Few geometric space-based has proposed in recent years has also gained popularity for finding communities. Mahmood and Small (2015) suggested a subspace-based community detection utilizing a sparse linear encoding using the least square method. The algorithm first computes the geodesic distance between every pair of nodes. After that, a similarity function is defined based on geodesic distance, and a similarity matrix is computed using similarity function and geodesic distance. Then a coefficient matrix is learned to represent a node(column) i as a linear combination of other nodes and store the linear coefficients and produce the linear encoding of the graph. High values of coefficients mean high similarity with that node(column), and lower values mean less similarity with the corresponding node. Finally, the algorithm applies spectral clustering to both adjacency matrix and linear encoding learned to find the communities. Ding et al. (2018) proposed a low-rank subspace learning-based community detection (LRSCD) algorithm. The algorithm performs low-rank decomposition to represent each node vector as a linear combination of the other nodes under the same subspace. After that, the proximity matrix is computed by the low-rank coefficient matrix, followed by spectral clustering to obtain the final communities. The above two methods learn low-dimensional vector representation or embedding features of nodes specific to community detection and then apply clustering algorithms to detect communities. In recent years, there are many remarkable network embedding methods have been devolved. In this paper, we use popular network embedding methods, first to find the low-dimensional vector representation of each node of the network followed by k-means clustering and then gravitational search algorithm to uncover the community structures. Hence, our proposed algorithm is a hybrid of geometric space-based and evolutionary-based methods.

3 Preliminaries

3.1 Network embedding

Many real-life complex systems are expressed as a graph such as online social networks, communication networks, biological networks. Conventional graph representation techniques may not be efficient in processing these voluminous network data to perform several analytic operations. Techniques that employ the representation of graph vertices in vector space so that different machine learning-based algorithms can be employed to perform various network analysis tasks. Such techniques are known as graph embedding or network embedding, which intend to learn a mapping of vertices to a low-dimensional space of

features and preserving the inherent properties and topological characteristics of the network. Network embedding methods intend to preserve first-order proximity, second-order proximity, and higher-order proximity to preserve the network structure, where first-order proximity refers to the similarity between two nodes if an edge connects them. Similarly, second-order proximity refers to the similarity of two nodes, which are at a 2-hop distance, and if they have a similar neighborhood. Recently, researchers have come up with many network embedding algorithms, and these methods can be broadly classified into three categories- factorization based, random walk based, deep learning-based (Goyal and Ferrara 2018; Cui et al. 2018). In this manuscript, we utilize one popular recent method for embedding from each category- HOPE (higher-order proximity preserved embedding) (Ou et al. 2016) from factorization based methods, node2vec (Grover and Leskovec 2016) from random walk based techniques, and SDNE (structural deep network embedding) (Wang et al. 2016) from deep learning-based category. Their brief descriptions are as follows:

(i) HOPE:- Higher Order Proximity Preserved Embedding (HOPE) is a Factorization based methods, which represent the connections between nodes in the form of a matrix and achieve the embedding by factorizing the matrix. HOPE intends to preserve asymmetric transitivity in case of a directed graph. The algorithm obtains a general formulation of a class of high-order proximity measurements like Katz index and then utilize generalized SVD to find the overall solution.

(ii) node2vec:- It is a random walk based technique that uses random walks from a vertex (u) and calculates the probability of reaching a vertex (v) in that random walk. The more a vertex (v) occurs in random walks together with another vertex, the more is the chance that vertices u and v are closely related. The different random walks are fed into a skip-gram neural network to generate embedding of the graph. Node2vec has two controlling parameters p and q. The idea is to make a trade-off between local and global views of a network. The parameter p defines how probable is that the random walk would return to the previous node whereas q defines how probable is that the random walk would discover the undiscovered part of the graph Hence, p control discovery of the microscopic view around the node, and q controls the discovery of the broader neighborhood.

(iii) SDNE:- SDNE (structural deep network embedding) is a semi-supervised deep learning-based network embedding method to preserve the neighbor structures of nodes in the network with a high non-linearity and sparse structure. The method uses multiple layers of non-linear functions to preserve the first-order proximity and second-order proximity. SDNE uses the notion of Laplacian eigenmaps to preserve first-order proximity of vertices that captures the local structure of the network. The second-order proximity is used by the unsupervised component to capture the global network structure.

Hence, the embedding technique captures the geometric structure of the graph, its vertex-to-vertex connection, and all other related information in the form of low-dimensional vector representation so that various machine learning-based algorithms can be employed to perform network analysis tasks. Also, there is a trade-off between graph properties considered and the size of vectors in embedding. The embedding capturing more features requires more dimension d, results in better accuracy, but is computationally expensive. We need to maintain the proper balance between the size or dimension (d) of the vectors and the computational cost.

3.2 Gravitational search algorithm

The gravitational search algorithm (GSA) is a nature-inspired optimization algorithm that tries to search a solution using rules similar to the gravitational laws of nature (Rashedi et al. 2009), which requires polynomial time with respect to the problem size. The laws of gravity govern the force between the particles, their acceleration, and velocity are used to search the solution. The basic steps of GSA is outlined in Algorithm 1. Random initial solutions are generated by taking into account the limits of the solution. The values in the solution represent their position in different dimensions. Then these solutions referred to as particles that are assigned a fitness value. The fitness value is the measure of how good our solution is. Mass is assigned to each solution (particles), and force is calculated between them according to the law of gravitational force. The force then leads to acceleration, and that causes a particle to acquire velocity.

Algorithm 1 Gravitational search algorithm.

Input: i : Search space dimensions,
S: Number of initial particles
Output: Optimal Solution
1: Select appropriate value for constants
2: $Particles \leftarrow$ generate_solution(i,S)
3: $Fitness \leftarrow$ calc_fitness(Particles)
4: $Masses \leftarrow$ compute_mass(Fitness)
5: $Acceleration \leftarrow$ calc_accelaration(Masses)
6: $Velocity \leftarrow$ compute_velocity(Acceleration)
7: $Particles \leftarrow$ get_new_position(Velocity,Particles)
8: Repeat Step-3 to 7 until stopping condition is met
9: Solution with best fitness value

Now, if we consider the time to be one unit, we can calculate the new position of the particles. We continue this until a terminating condition is met. The particles with better fitness will tend to attract the particles with lower fitness values towards themselves, so the fitness values of all the particles will improve after many iterations. GSA depends heavily on the choice of the initial population. A randomly chosen initial population may get struck into local minima and does not produce a near-optimum result. Usually, In GSA the initialization is random and hence take more iterations and computations to reach close to the optimum solution. If proper input is chosen, then GSA explores the solution space properly and provides an optimum result. The k-means result provides a good initial population to GSA. Hence, with better exploration near the result of k-means, GSA can provide a high-quality solution in fewer iterations. GSA has been utilized in many engineering applications like combinatorial optimization problems, clustering, and classification problem, training neural networks (Siddiquea and Adelib 2016).

4 Proposed community detection algorithm

This section presents the details of the algorithm adopted in this paper, which utilizes network embedding technique and gravitational search algorithm (GSA) to identify the communities in the network. These network embedding techniques aims to learn

low-dimensional representations of vertices in networks, intending to capture and preserve the network structure. Using network embedding techniques, we consider nodes as points in d- dimensional vector space, such that two nodes that are connected by an edge are closer in feature space and non-connected nodes are away. The network embedding techniques aim to learn low-dimensional representations of vertices in networks and capturing the network structure. Using network embedding techniques, we consider nodes as points in d- dimensional vector space, such that two nodes that are connected by an edge are closer in feature space, and non-connected nodes are away. We intend to exploit such low-dimensional vector representation of nodes and the structural features of the network to perform community detection. The outline of the proposed algorithm is presented in Algorithm 2.

Algorithm 2 Proposed community detection algorithm.

Input: $G = (V, E)$: Graph, k : Number of communities , $true_labels$
Output: NMI

$d \leftarrow$ select an appropriate dimension for embedding
$method_name \leftarrow$ chose a method for embedding
$emb \leftarrow$learn_embedding($G, d, method_name$)
$var \leftarrow$ Choose appropriate var. for low rank approx.
$emb_low \leftarrow$ low_rank_approximation (emb, var)
$predicted \leftarrow$ k-means (emb_low, k)
$gsa_result \leftarrow$ GSA (predicted, S, r_i, r_j, iter,G, α)
$NMI \leftarrow$ calc_NMI (gsa_result, truth_labels)
$Modularity \leftarrow$ calc_modularity (gsa_result, truth_labels)
$Precision \leftarrow$ calc_precision (gsa_result, truth_labels)
$Recall \leftarrow$ calc_recall (gsa_result, truth_labels)
return $NMI, Modularity, Precision, Recall$

The detailed descriptions of the proposed algorithm are as follows:

- Step 1: We first choose the appropriate vector space size, d, to embed the nodes of the graph into low-level vector representations. If we choose too small d, then embedding may not be able to capture the complete information from the graph. The more the dimension, the better it can achieve graph properties, but high computation time it requires. Here, d is a hyper-parameter, higher the value d captures the more graph properties. In the case of a simple graph with a uniform distribution of edges, even smaller value of d can properly capture the graph properties. If the graph is more complex, then we should select the d by considering the trade-off between graph properties to capture and computational complexity.
- Step 2: There are three classes of embedding methods, namely graph factorization based, random walk based, and deep learning-based. We consider one latest technique from each category to compare the results of community detection. We use the HOPE method from the graph factorization category, node2vec method from the random walk category, and SDNE method from the deep learning category.
- Step 3: In this step, we generate the vector embedding from the input graph (G), dimension (d), and the chosen method for embedding either HOPE, node2vec, or SDNE. The outcome of this step is the nodes of the input graph, each represented in a d dimensional vector space that captures the features of the network. Here, emb is a

two-dimensional matrix containing all nodes along with their d dimensional low-level vector representations.

- Step 4: To avoid the results from being affected by noise or outliers, we remove noise by low-rank approximation of the learned embedding. For low-rank approximation, we need to choose the amount of variance to preserve, which is also a hyperparameter. Therefore, this step involves the selection of variance for low-rank approximation. The variance should be optimum to remove unwanted noise.
- Step 5: In this step, we learn low-rank approximation using SVD (Singular Value Decomposition) method on embedded matrix emb as obtained in step 3 with chosen variance in the previous step. The diagonal matrix of SVD represents variance along different dimensions. Depending on the amount of variance to preserve, we keep some diagonal values and make others as zero. The result obtained in this step is stored in the new matrix, emb_low.
- Step 6: In this step, we apply k-means clustering algorithm to the result of low-rank approximation obtained in the previous step along with the number of clusters as input. We employ k-means technique as it is simple and scales well to large data items. Also, k-means finds a near-optimal solution in a feasible amount of time. We use k-means++ initialization for better results. The result of the k-means is a $k \times d$ dimension vector representing coordinates of k cluster centers.
- Step 7: The performance of k-means depends on the initial position of centroids and may confine in local optima. The gravitational search algorithm (GSA) is a famous population optimization-based technique based on the law of gravity. GSA is used to obtain a near-optimal solution and hence is a potential way to improve the performance of k-means. In this step, we perform a GSA search with an initial population of S particles, and the number of iterations (iter). It also takes parameters initial gravitational constant (G), the decay rate of the gravitational constant (α), and two random coefficients r_i and r_j for force and velocity equation in GSA respectively.

  Note that community detection is an NP-hard problem and an optimal solution can not be obtained in polynomial time unless $P = NP$. As it is unlikely that $P = NP$, one needs to find a near optimal solution in polynomial time. This is what GSA tries to do. But it is expected that if the results of k-means are input to GSA, high chances are there that GSA will be able to find a near optimal solution. Since, GSA is based on both exploration and exploitation, even if k-means result is struck in local optima, GSA will explore other solutions too and may lead to a better solution, which is more likely to be a global optimum or near global opimum. The execution and efficiency of the GSA algorithm mainly depend on S (number of initial solution particles) and iter (number of iterations). The value of S should be kept reasonable with respect to the problem size. A smaller value may not provide a good exploration whereas a large value will make the algorithm computationally expensive and may increase noise in search space. The outcome of this step is a two-dimensional vector having coordinates of k clusters flattened into a $k \times d$ dimension space. The outcome of this step is a two-dimensional vector having coordinates of k clusters flattened into a $k \times d$ dimension space and hence the result of the community detection of the input graph. We made the following changes to the original gravitational search algorithm, as mentioned in Algorithm 2, to adapt in our proposed approach and leading to optimize results in community detection:-

  (a) Initial Population: The initial population is not random, but it has the result of k-means clustering as one of the inputs. The rest of the particles are generated randomly.

(b) Fitness Function: The calc_fitness function, as in line number 3 of Algorithm 1 is defined as the squared sum of euclidean distances between the cluster centers in solution particle and the nodes belonging to that cluster. When cluster centers are close to all the points belonging to the cluster, the fitness value may be small. The optimization is a minimization problem to minimize this fitness value.

(c) Terminating Condition:- Since we are improving the results of clustering through GSA, stating whether we achieved minimum fitness is difficult. Hence, we run the GSA algorithm for a certain fixed number of iterations, and the solution with the best fitness value is returned as a result.

– Step 8 − 11: For effective evaluation of the obtained results, we analyze different performance criteria. Normalized Mutual Information (NMI) (McDaid et al. 2013) is an entropy-based evaluation criteria. To calculate NMI, we need to first calculate the community entropy with ground truth $H(Y)$ and predicted labels $H(C)$. Then we calculate the conditional probabilities, and then NMI is calculated using a series of equations from (1) to (6). To calculate the modularity, which is a measure of division among communities, is calculated using (7). Precision and recall are a measure of relevance of the result obtained and are calculated using (8) and (9), respectively.

We present the complete simulation of the proposed algorithm on Zachary Karate Club dataset having two ground-truth communities using embedding in the vector space of a minimal dimension (d)= 2, as depicted in Fig. 1. Figure 1a shows the karate dataset as input with 34 nodes and 78 edges. The input network is then converted to 2D coordinates using the node2vec embedding technique. In Fig. 1b each point represent a node in the

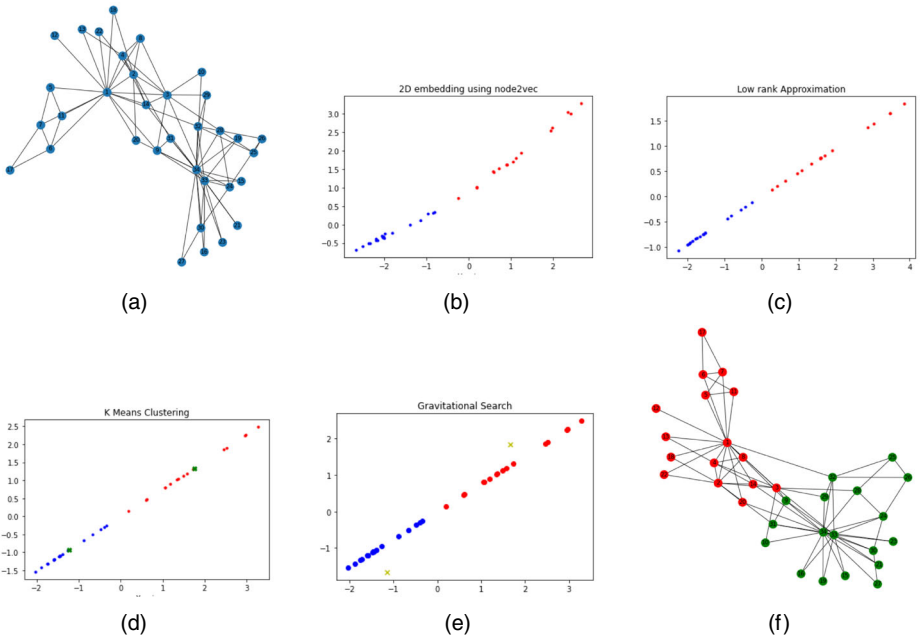


Fig. 1 Simulation of proposed Algorithm on Zachary Karate Club network with $2-d$ node2vec embedding-**a** input Zachary Karate Club dataset, **b** 2-D embedding using node2vec, **c** Low Rank Approximation, **d** Cluster Centers With k-Means, **e** Cluster Centers after GSA, and **f** Finally, obtained communities as output

network, and the color of each coordinate point shows the community to which it belongs according to the ground truth value. Figure 1c shows the new coordinates of the points after performing low-rank approximation and removing noise. Notice that the points are almost in a straight line making it easier to define boundaries of the clusters. Figure 1d depicts two cluster centers obtained by applying k-means. The result of k-means is optimized through the GSA, and new and improved cluster centers are obtained through GSA as presented in Fig. 1e. The cluster centers seem to be far apart from coordinates points but are better at classifying each coordinate into a different community. The final network divided into two communities, with red and green colors, each color representing different communities, as shown in Fig. 1f.

5 Datasets and evaluation criteria

For a better evaluation of the proposed algorithm, we choose different real-life and synthetic graph datasets with varying sizes, degrees of intermixing, density, and the number of clusters. We adopt eight datasets, out of which five are real-world datasets, and rest three are synthetic. Karate Club (Zachary 1977), American Football (Girvan and Newman 2002), Polblogs (Adamic and Glance 2005), Polbooks (Krebs 2020), Word (Newman 2006b) are the real-world networks datasets and three synthetic datasets created using LFR benchmark (Lancichinetti et al. 2008) with 128 nodes and varying mixing parameter (μ). For each chosen dataset, the ground-truth community value is known for each vertex. The brief description of each dataset is as follows:

5.1 Real-life datasets

(i) Zachary Karate Club (Zachary 1977): A social network of 34 members of a karate club at a US university in the 1970s. The nodes represent the members of the club and edges represent the friendship relationship between them. The network with 34 nodes has 78 edges and average degree 4.88. It is one of the most popular dataset used in community detection with two ground truth communities.

(ii) American Football (Girvan and Newman 2002): It is a network of football games between Division IA colleges during the Fall season, 2000. The network has 115 nodes and 613 edges with an average degree of 10.66, and 12 predefined communities.

(iii) Political Blogs (Adamic and Glance 2005): A network of hyperlinks between different weblogs on US politics, recorded in 2005 by Adamic and Glance. The political blogs network with 1490 nodes, 16718 edges, and an average degree of 25 with two ground-truth communities.

(iv) Political Books (Krebs 2020): A network of books on US politics published around the time of the 2004 election for president. Nodes represent books, and edges between books represent frequent co-purchasing of books by the same buyers. This network has 105 nodes and 441 edges with an average degree of 8. This is a relatively medium-sized dataset with three ground-truth communities.

(v) Word (Newman 2006b): It is a network of common words (nouns and adjectives) in the novel David Copperfield by Charles Dickens. This network has 112 nodes, and 425 edges with an average degree of 7. This is a relatively medium-sized dataset with loosely defined communities and 2 ground-truth communities.

5.2 Synthetic datasets

The datasets in this section are created using the LFR benchmark for an undirected graph (Lancichinetti et al. 2008). The create three datasets having all characteristics similar, but the level of intermixing among communities is varied. The level of intermixing is defined by the mixing parameter, which is the average fraction of neighboring nodes of a node that do not belong to any community that the benchmark node belongs to. At the extremes, when the parameter is set to 0, all links are within community links, and for mixing parameter=1, all links are between nodes belonging to different communities. We use three different variant of LFR datasets.

(i) LFR-1: It is a synthetic dataset created using LFR benchmark for undirected graphs. The graph has 128 nodes and an average degree of 16. The other LFR parameters are community size as 32, mixing parameter (μ) as 0.1, and exponent for weight distribution as 1. The graph has four ground-truth communities and is a graph with dense communities and less inter-mixing.

(ii) LFR-2:- It is a synthetic dataset created using the same LFR benchmark. All parameters are same as of LFR-1 except the mixing parameter 0.3. The graph has 128 nodes and average degree 16. The graph has 4 ground truth communities and is a graph with dense communities and small amount of inter-mixing.

(iii) LFR-3:- It is also a synthetic dataset created using the same LFR benchmark. All parameters are same as of LFR-1 except the mixing parameter which is set to 0.5 to get a graph with more intermixing among communities. The graph has 128 nodes and average degree 16 with four ground truth communities.

5.3 Evaluation criteria

(i) Normalized Mutual Information: The normalized mutual information (NMI) (Estévez et al. 2009) of two random variables is a normalized measure of the mutual dependence relation between the two variables. More specifically, it quantifies the "amount of information" obtained about one random variable by observing the other random variable. The concept of mutual information is intricately linked to that of entropy of a random variable, a fundamental notion in information theory that quantifies the expected "amount of information" held in a random variable. The three steps involved in the calculation are:

Calculating Entropy of Community Labels $H(Y)$ represents the entropy of communities using ground truth, whereas $H(C)$ represent the entropy using the predicted labels. Here n represents the total number of distinct communities in the ground truth, and m represents the number of distinct communities in the ground truth. Similarly, G_i represents probability of community i using ground truth whereas P_i represents probability calculated using predicted labels.

$$H(Y) = \sum_{i=1}^{n} G_i log(G_i) \tag{1}$$

, where G_i is the number of nodes belonging to community i using ground truth labels up on total number of nodes in the network.

$$H(C) = \sum_{i=1}^{m} P_i log(P_i) \tag{2}$$

, where P_i is the number of nodes belonging to community i using predicted labels up on total number of nodes in the network.

Calculating Conditional Entropy ($H(Y|C)$):- This is dependent on the outcome of the algorithm. It is calculated for each community C, and its sum is used further. Here n represents the total number of clusters in which the nodes are divided.

$$H(Y|C) = \sum_{i=1}^{n} P(i) \sum_{j=1}^{m} f(i, j) \tag{3}$$

$$f(i, j) = P(Y = j|C = i) log(P(Y = j|C = i)) \tag{4}$$

Calculating NMI:- Before calculating NMI, we calculate MI (Mutual Information). Then NMI is calculated by normalizing MI using (6).

$$MI = H(Y) - H(Y|C) \tag{5}$$

$$NMI = \frac{2 * MI}{H(Y) + H(C)} \tag{6}$$

(ii) Modularity: Modularity (Q) is designed to measure the strength of the division of a network into modules or communities. Modularity is often used for evaluating the qualities of the detecting community structure in networks. The value of modularity is computed using (7).

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \tag{7}$$

where, m represents the number of edges in the network, A is the Adjacency matrix of the network, k_i is the degree of i and $\delta(c_i, c_j)$ is 1 if i and j are in the same community and 0 otherwise. The term $k_i k_j$ represents the probability that a random edge would go between i and j.

(iii) Precision and recall: precision, also called positive predictive value, is the fraction of relevant instances among the retrieved instances, while recall, also known as sensitivity, is the fraction of the total amount of relevant instances that were actually retrieved. Both precision and recall are, therefore, based on an understanding and measure of relevance. Precision shows the agreement between the known labels and predicted labels, whereas recall shows the effectiveness of the algorithm in predicting labels.

$$Precision = \frac{\sum_{i=1}^{n} tp_i}{\sum_{i=1}^{n} (tp_i + fp_i)} \tag{8}$$

$$Recall = \frac{\sum_{i=1}^{n} tp_i}{\sum_{i=1}^{n} (tp_i + fn_i)} \tag{9}$$

, where tp_i, fp_i, fn_i represents true positives, false positives and false negative for community c_i and n represents the total number of communities.

6 Experimental results and analysis

To properly evaluate the proposed algorithm, we chose a variety of networks varying in size, complexity, and number of communities. We perform the investigations on each of the selected real-life and synthetic datasets, as mentioned in the Sections 5.1, and 5.2. To perform the low-rank approximation using SVD, the variance is set to 50 to extract important information and leave out the less useful information. The initial population for GSA is set

to 20, as we got the optimal performance in some datasets on this particular value of the initial population. The performance of the GSA depends on the initial population for different dimensions, to maintain uniformity in evaluation for all datasets, we chose the same the initial population. The rate of decay of gravitational constant determines the rate at which the force will decay, a higher value may lead to a local optimum, whereas a smaller value may not provide result in a reasonable number of iterations. To maintain a balance, the value of the decay rate of the gravitational constant is set to 20. We fix the number of iterations for the gravitational search algorithm as 100 as it is a balance value to obtain good results in a reasonable time. We obtain the predicted labels for each node with each of the three embedding methods, namely HOPE, node2vec, and SDNE. We perform the embedding till 128 dimensions of the low-dimensional vector representation. The results are evaluated based on the four performance measures, namely NMI, modularity, precision, and recall, as listed in Section 5.3. Figure 2 depicts the pictorial view of results of community detection on various datasets. From Fig. 2a to f obtained results are compiled using node2vec embedding technique except in Fig. 2d for Word dataset result is drawn using SDNE.

6.1 Normalized mutual information

Normalized Mutual Information (NMI) is a normalization of the mutual information (MI) score to scale the results between 0 and 1, where 0 corresponds to no mutual information, and 1 corresponds to perfect correlation. We observe a general trend is that NMI value increases with the increase in the dimension of embedding. Node2vec gives the best result except in the case of sparse graphs like Word, where SDNE performs better. Also, HOPE performs better than SDNE with datasets like American Football having more number of communities (ref. Fig. 3b), but with a lesser number of communities like Word, SDNE dominates, which is evident from Fig. 3e. Both are approximately the same sized datasets, but American Football has more communities than Word. The gravitational search help in finding better results in lower dimensions, so we see more flat sections in the plot and better values even at lower dimensions. GSA plays a significant role in large networks as it leads to better-separated communities with even with low dimension embedding. Reducing dimensions in embedding saves computational time as it is the most time-consuming subprocess in community detection. The brief description of the NMI results obtained on each dataset is discussed below:-

Zachary Karate Club (Fig. 3a):- Node2vec is performing best, with NMI value 1 for dimension greater than 32. Hope is better than SDNE for lower dimensions but saturates sooner, and for higher dimension value for SDNE and HOPE are almost equal.

American Football (Fig. 3b):- For smaller dimensions, the NMI values are less for all three embedding methods. However, it increases as we increase the dimension and then remains constant for dimension greater than 32. Node2vec gives the best result, followed by HOPE and then SDNE. In this case, HOPE performs better then SDNE.

Political Blogs (Fig. 3c):- It is a relatively large dataset with a lesser number of communities. Here, node2vec gives a throughout better result when compared with other methods. SDNE performs well as compared to HOPE with lower dimensions. For higher dimensions, NMI values with HOPE improves but slightly less than that of SDNE, and NMI values are constant for dimension greater than 16 for all datasets. For larger datasets, the algorithm tends to give better results even at low dimensions as GSA improves the clustering.

Political Books (Fig. 3d):- Here also, node2vec gives a better result even for smaller dimensions of embeddings, but for higher dimensions, SDNE gives NMI values close to that of node2vec. The result remains unchanged for a dimension greater than 32.

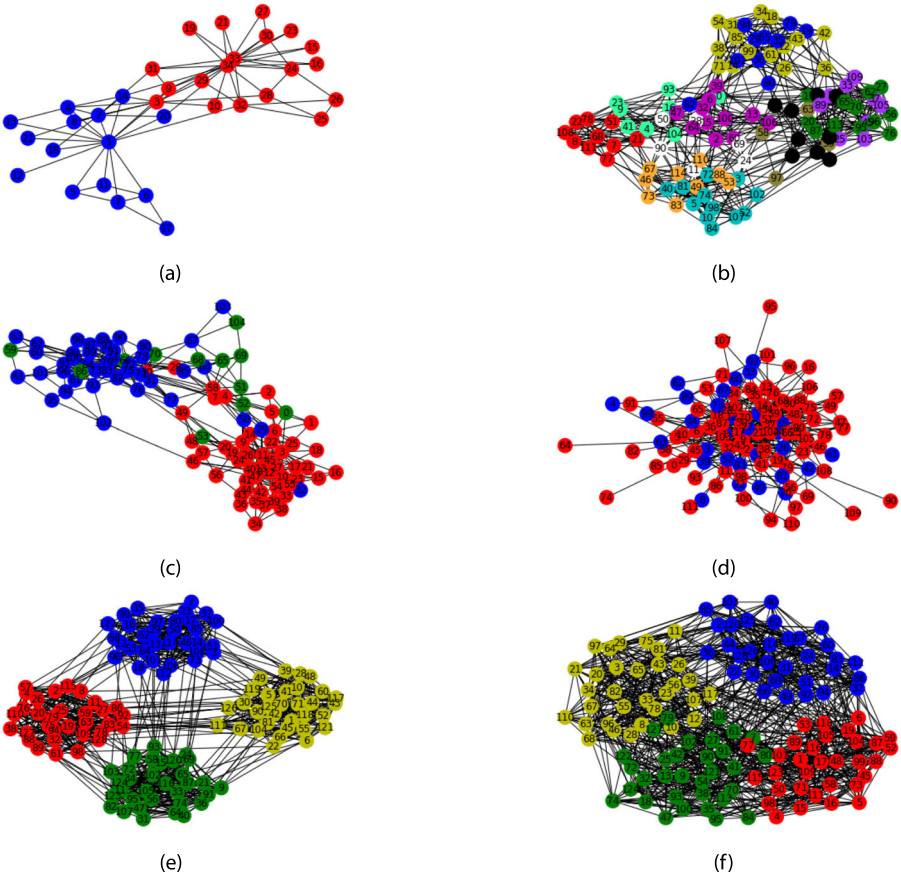


Fig. 2 Community detection results on various networks **a** Karate club with two visible communities, **b** American Football with twelve visible communities, **c** political Books with three communities, **d** Word dataset with two communities, **e** LFR-1, a synthetically generated dataset, and the graph is well separated into four communities, **f** community detection on LFR-2 dataset the division among the four communities is visible

Word (Fig. 3e):- In the case of the Word dataset, SDNE performs better for all dimensions. Node2vec, which produces the best results in other graphs, fails to give significant results for this dataset due to not so well defines communities.

LFR-1 (Fig. 3f):- In a graph with less intermixing among communities, HOPE and node2vec show the same trend. The NMI values of both node2vec and HOPE achieves optimum with an embedding dimension of 16 due to GSA. SDNE gives a poor result with fewer dimensions but improves as we increase dimension.

LFR-2 (Fig. 3g):- The trends captured in LFR-1 are almost the same in LFR-2 for node2vec and sdne, both giving maximum value for embedding dimension greater than 16. Here, the performance of SDNE does not improve well with a higher dimension.

LFR-3 (Fig. 3h):- Though the trends remain the same in LFR-3, maximum obtained NMI values decrease significantly due to more intermixing among communities. Node2vec performs better if the intermixing is more. After the embedding dimension of 32, performance remains the same for all the methods of embedding.

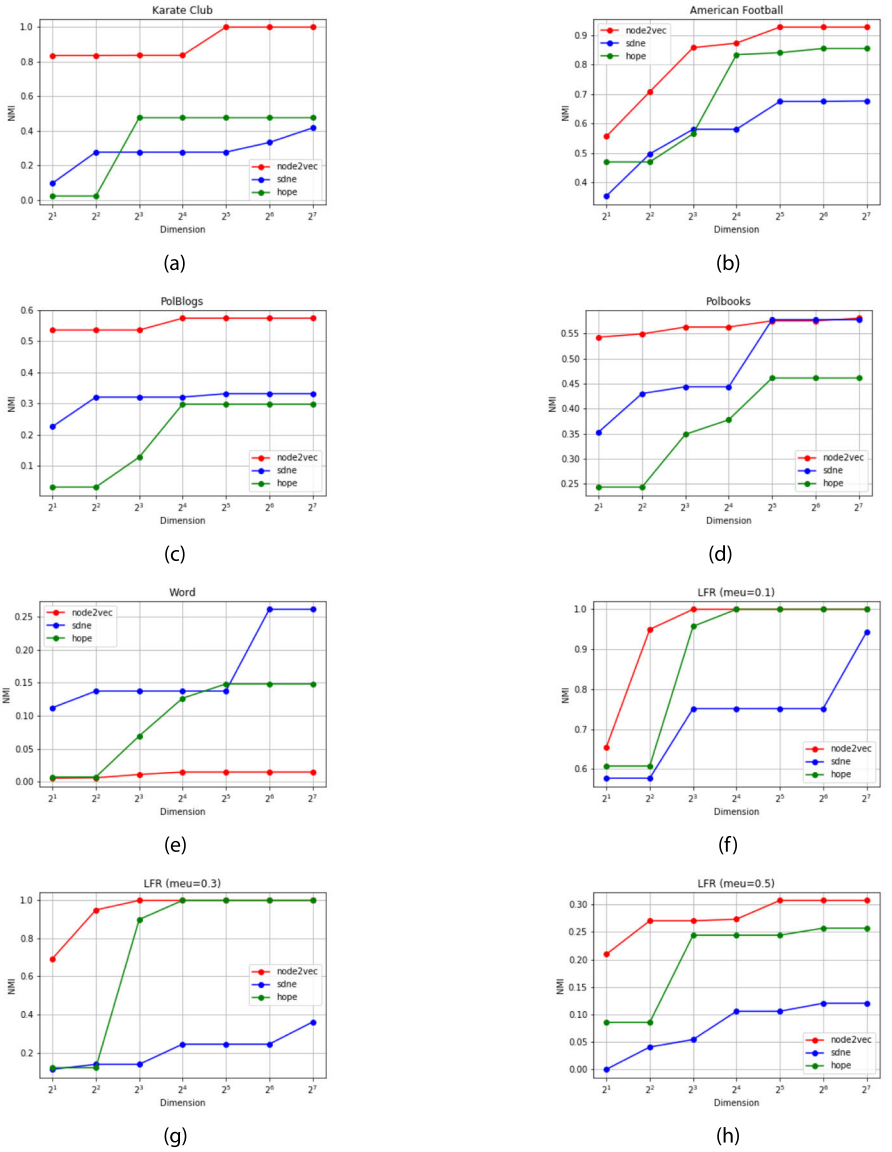


Fig. 3 NMI results on different datasets **a** Karate club, **b** American Football, **c** Polblogs , **d** political Books, **e** Word dataset **f** LFR-1, **g** LFR-2, and **h** LFR-3

We compare the performance of our proposed algorithm namely network embedding using HOPE (NE-HOPE), network embedding using SDNE (NE-SDNE) and network embedding using node2vec (NE-N2V) with some well-known approaches of community detection like Girvan-Newman (Edge-Betweenness) method (Girvan and Newman 2002), fast greedy (Newman 2006a), label propagation (Cordasco and Gargano 2010), and infomap (Rosvall and Bergstrom 2008). Table 1 lists the value of normalized mutual information (NMI) obtained on various datasets by the well-known approaches of community detection

Table 1 NMI results of different methods on various used datasets where FG, LP, GN, NE-HOPE, NE-SDNE, NE-N2V refers to fast greedy, label propagation, Girvan-Newman, network embedding using HOPE, network embedding using SDNE, and network embedding using node2vec methods respectively

Dataset	Infomap	FG	LP	GN	NE-HOPE	NE-SDNE	NE-N2V
Karate	0.8371	0.6924	0.1733	0.8364	0.4764	0.3333	**1**
Football	0.8829	0.6977	0.1193	0.3592	0.8543	0.6745	**0.9268**
Word	0.0332	0.0052	0.0836	0.0030	0.1480	**0.2610**	0.0147
Polbooks	0.5368	0.5308	0.1419	0.4051	0.46128	**0.57764**	0.5752
Polblogs	0.3363	0.37824	0.0019	0.1161	0.2972	0.3311	**0.5732**
LFR-1	**1**	**1**	0.1072	0.5771	**1**	0.7515	**1**
LFR-2	**1**	0.9497	0.0960	0.5771	**1**	0.2440	**1**
LFR-3	0.0064	**0.4233**	0.1096	0.0152	0.2571	0.1203	0.30764

along with the proposed approach with three variants by taking the size of embedding (d) as 64. From the results, it is evident that our network embedding based algorithms perform well.

6.2 Modularity

Modularity is a popular measure to obtain the strength of the division of a network into modules. We observe an almost similar trend in modularity, like that of NMI for all the datasets. Node2vec is the best technique that gives the most modular division of graph among communities. Employing GSA shows better results even at lower dimensions.

Karate Club (Fig. 4a):-In this dataset, node2vec gives a more modular community division than other embedding techniques, which is similar to NMI-dimension results, as in Fig. 3a. The performance of HOPE remains after node2vec, and SDNE performs least. There is also no improvement for dimensions greater than 8 in each method.

American Football (Fig. 4b):- In terms of modularity, node2vec is the best among three embeddings. Here, SDNE gives a better result than HOPE for higher dimensions, which is the opposite of what we observe from the NMI-dimension graph, as in Fig. 3b. Here, the performance of SDNE and HOPE are similar for the embedding dimension 32 onwards.

Political Books (Fig. 4c):- The trend in modularity value is almost the same as compared to NMI, but node2vec tends to give a more modular distribution than SDNE, although the NMI values were almost the same with node2vec and SDNE for dimension greater than 32.

Political Blogs (Fig. 4d):- This relatively large dataset shows a near constant plot with node2vec. SDNE does not perform good for low dimesions but the modularity value is near to node2vec values when dimension is kept higher.

Word (Fig. 4e) Word dataset is a relatively sparse graph. The modularity trend in this dataset changes entirely in comparison to NMI. In NMI results, SDNE performed best. However, here in modularity, node2vec produces better than both SDNE and HOPE. For HOPE, there is no change seen in modularity even we increase the dimension of embedding. Also, SDNE performs inadequately in this case.

LFR-1 (Fig. 4f):- In a graph with less inter-mixing among communities, modularity, and NMI, both metrics tend to show the same trend with node2vec giving better results for lower dimensions. SDNE and HOPE also perform better for larger dimensions.

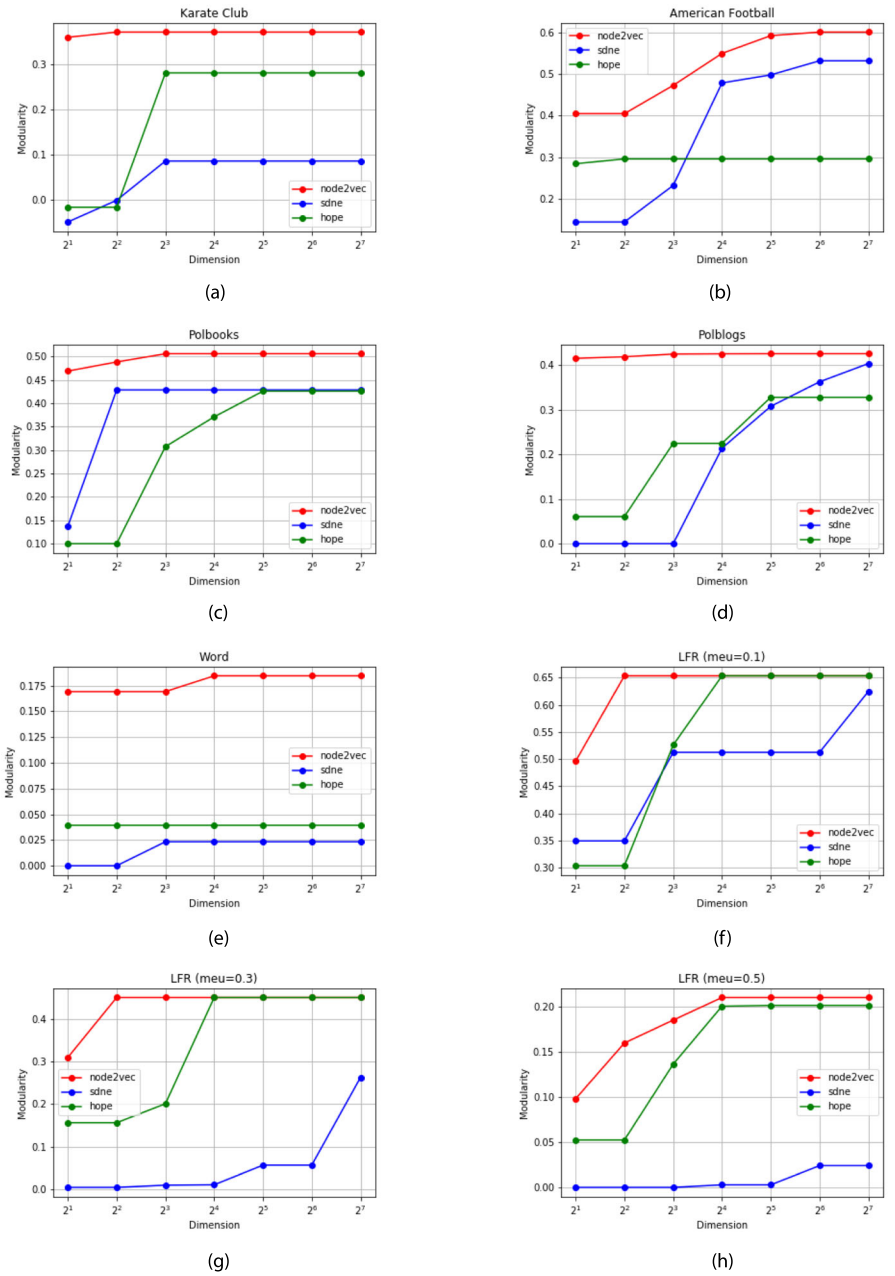


Fig. 4 Modularity results on different datasets **a** Karate club, **b** American Football, **c** political Books, **d** Polblogs **e** Word dataset, **f** LFR-1, **g** LFR-2, and **h** LFR-3

LFR-2 (Fig. 4g):- For a lower dimension of embedding, node2vec gives better results. However, for dimension 32 onwards, the performance of node2vec and HOPE are similar. The performance with SDNE is not good in terms of modularity when intermixing is increased.

Table 2 Modularity results of different methods on various used datasets where FG, LP, GN, NE-HOPE, NE-SDNE, NE-N2V refers to fast greedy, label propagation, Girvan-Newman, network embedding using HOPE, network embedding using SDNE, and network embedding using node2vec methods respectively

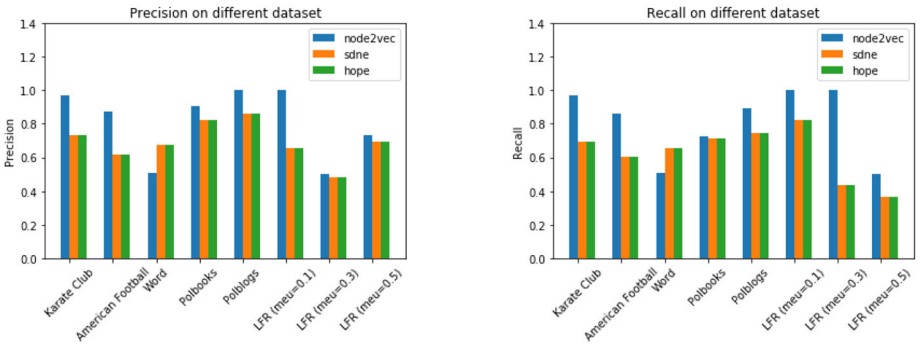
Dataset	Infomap	Greedy	LP	Girwan	NE-HOPE	NE-SDNE	NE-N2V
Karate	**0.4020**	0.3806	0.3547	0.3599	0.2813	0.0864	0.3714
Football	0.6008	0.5497	0.5509	0.4003	0.2957	0.5319	**0.6010**
Word	0.0084	**0.2946**	0.0074	0.0092	0.0395	0.0233	0.1845
Polbooks	0.5267	0.5019	0.4811	0.4428	0.4258	0.4284	**0.5357**
Polblogs	0.4248	**0.4270**	0.4262	0.0012	0.3275	0.3621	0.4249
LFR-1	**0.6533**	**0.6533**	**0.6533**	0.3281	**0.6533**	0.5123	**0.6533**
LFR-2	**0.4521**	0.4361	0.2645	0.2246	**0.4521**	0.0560	**0.4521**
LFR-3	0.0019	0.2088	0.0011	0.0001	0.2016	0.0241	**0.2104**

LFR-3 (Fig. 4h):- As communities in this graph dataset tend to get more inter-mixed, node2vec performs slightly better than HOPE. The results with the SDNE technique deteriorates as intermixing is increased.

We examine our network embedding based proposed algorithm based on modularity value with some well-known approaches of community detection like the Girvan-Newman (Edge-Betweenness) method (Girvan and Newman 2002), fast greedy (Newman 2006a), label propagation (Cordasco and Gargano 2010), and infomap (Rosvall and Bergstrom 2008). Table 2 presents the modularity value received on different datasets by various community detection methods considered and the proposed approach using network embedding by taking the size of embedding (d) as 64. From the results, it is clear that network embedding using node2vec (NE-N2V) algorithm outperforms others in most of the cases.

6.3 Precision

Precision is the fraction of relevant instances among the retrieved instances. Figure 5a depicts the precision obtained on the various dataset with different embedding techniques. Precision value is best in the node2vec embedding technique for all the datasets except in case of sparse graphs like Word. For Karate club, LFR-1, and LFR-2 datasets, the value of



(a) Precision (b) Recall

Fig. 5 Precesion and recall value of different methods in each dataset

precision obtained is almost one using the node2vec embedding technique. However, the precision value of SDNE and HOPE is almost the same on all datasets. This implies that although SDNE and HOPE predict the labels with the same accuracy, but they vary in terms of structural division and having different modularity and NMI on different datasets.

6.4 Recall

Figure 5b presents the recall value obtained on various datasets. The recall trends are the same as precision Section 6.3, where node2vec performs better on all the datasets except Word. Here, node2vec tends to perform even better in terms of recall on all the datasets except Word. For Karate club, LFR-1, and LFR-2 datasets, the value of precision achieved is almost one using the node2vec embedding technique. This indicates that node2vec is a more sensitive embedding technique and is good at predicting relevant instances.

7 Conclusion

Community detection is one of the most studied problems in the field of network science. In this paper, we utilized the idea of network embedding and gravitational search algorithm to uncover community structure in complex networks. We obtained the low-dimensional vector representation of nodes by three popular methods of embedding techniques i.e, HOPE from matrix factorization category, node2vec from random walk, and SDNE from deep learning category. Then we apply low-rank approximation followed by k-means clustering. Then gravitational search algorithm (GSA) is used to improve the results of centroids of clusters obtained through k-means leading to better results even at low dimensions. As the output solution of the k-means clustering is one of the member of initial population of GSA, our algorithm does not trap in local optima, which usually happens with the evolutionary-based algorithm and clustering techniques like k-means. We performed comprehensive investigations on five real-life, and three synthetic network datasets using various performance measures, and the results obtained are satisfactory enough to suggest that the proposed algorithm presents a general framework for detecting communities in a variety of networks using network embedding techniques. Comparing benchmarks of various traditional embedding techniques, we achieved the best results with node2vec embedding for all the graphs except networks where communities have a high level of intermixing, SDNE performs better in such a scenario. This is due to how these embedding techniques operate. Node2vec is a random walk based embedding technique and is better at capturing path based relationships between the nodes, so it is proved to be more a modular embedding technique. The node2vec performs better if community boundaries are simpler. SDNE, on the other hand, is good at capturing complex relationships between nodes, which is evident from results on dataset like Word. In precision and recall criteria, SDNE and HOPE perform almost similar but node2vec performs better than others for all the datasets except on the Word dataset. Our study is a noble effort to employ popular network embedding methods for the community detection problem, and try to span the gap between machine learning techniques, evolutionary approaches, and network analysis operation. The extension of our research work can be to find a better fitness function to get better results using GSA. Different statistical distributions can also be used to generate the initial population and study their effect on the results.

References

Adamic, L.A., & Glance, N. (2005). The political blogosphere and the 2004 US Election. In *Proceedings of the WWW-2005 workshop on the weblogging ecosystem*.

Aggarwal, C.C. (2011). An introduction to social network data analytics. In *Social network data analytics* (pp. 1–15). Boston: Springer.

Aggarwal, C., & Subbian, K. (2014). Evolutionary network analysis a survey. *ACM Computing Surveys (CSUR)*, *1;47*(1), 1–36.

Arasteh, M., & Alizadeh, S. (2019). A fast divisive community detection algorithm based on edge degree betweenness centrality. *Applied Intelligence*, *49*(2), 689–702.

Bamakan, S.M., Nurgaliev, I., Qu, Q. (2109). Opinion leader detection: a methodological review. *Expert Systems with Applications*, *115*, 200–22.

Battiston, F., Iacovacci, J., Nicosia, V., Bianconi, G., Latora, V. (2016). Emergence of multiplex communities in collaboration networks. *PLoS ONE*, *11*(1), e0147451.

Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), P10008.

Chouchani, N., & Abed, M. (2020). Online social network analysis: detection of communities of interest. *Journal of Intelligent Information Systems*, *54*(1), 5–21.

Clauset, A., Newman, M.E., Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6), 06611.

Cordasco, G., & Gargano, L. (2010). Community detection via semi-synchronous label propagation algorithms. In *2010 IEEE international workshop on: business applications of social network analysis (BASNA)* (pp. 1–8): IEEE.

Cui, P., Wang, X., Pei, J., Zhu, W. (2018). A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering*, *31*(5), 833–52.

Danon, L., Arenas, A., Díaz-Guilera, A. (2008). Impact of community structure on information transfer. *Physical Review E*, *3;77*(3), 036103.

Ding, Z., Zhang, X., Sun, D., Luo, B. (2018). Low-rank subspace learning based network community detection. *Knowledge-Based Systems*, *155*, 71–82.

Ding, X., Zhang, J., Yang, J. (2018). A robust two-stage algorithm for local community detection. *Knowledge-Based Systems*, *152*, 188–199.

Estévez, P.A., Tesmer, M., Perez, C.A., Zurada, J.M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, *20*(2), 189–201.

Eustace, J., Wang, X., Li, J. (2014). Approximating web communities using subspace decomposition. *Knowledge-Based Systems*, *70*, 118–127.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, *486*(3), 75–174.

Fortunato, S., & Hric, D. (2016). Community detection in networks: a user guide. *Physics Reports*, *659*, 1–44.

Freeman, L. (2004). The development of social network analysis. *A Study in the Sociology of Science*, *1*, 687.

Girvan, M., & Newman, M.E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, *11;99*(12), 7821–6.

Goyal, P., & Ferrara, E. (2018). Graph embedding techniques, applications, and performance: a survey. *Knowledge-Based Systems*, *151*, 78–94.

Grover, A., & Leskovec, J. (2016). node2vec: scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 855–864): ACM.

Guerrero, M., Montoya, F.G., Baños, R., Alcayde, A., Gil, C. (2017). Adaptive community detection in complex networks using genetic algorithms. *Neurocomputing*, *266*, 101–113.

Gui, Q., Deng, R., Xue, P., Cheng, X. (2018). A community discovery algorithm based on boundary nodes and label propagation. *Pattern Recognition Letters*, *109*, 103–9.

Guo, K., He, L., Chen, Y., Guo, W., Zheng, J. (2020). A local community detection algorithm based on internal force between nodes. *Applied Intelligence*, *50*(2), 328–40.

Honghao, C., Zuren, F., Zhigang, R. (2013). Community detection using ant colony optimization. In *2013 IEEE congress on evolutionary computation* (pp. 3072–3078): IEEE.

Huang, H., Shen, H., Meng, Z., Chang, H., He, H. (2019). Community-based influence maximization for viral marketing. *Applied Intelligence*, *49*(6), 2137–50.

Jaradat, A.S., & Hamad, S.B. (2018). Community structure detection using firefly algorithm. *International Journal of Applied Metaheuristic Computing (IJAMC)*, *9*(4), 52–70.

Krebs, V. (2020). unpublished, http://www.orgnet.com/.

Kumar, S., & Panda, B.S. (2020). Identifying influential nodes in social networks: neighborhood coreness based voting approach. *Physica A: Statistical Mechanics and its Applications*, 124215.
Lancichinetti, A., Fortunato, S., Radicchi, F. (2008). Benchmark graphs for testing community detection algorithms. *Physical review E*, *78*(4), 046110.
Ma, L., Huang, H., He, Q., Chiew, K., Liu, Z. (2014). Toward seed-insensitive solutions to local community detection. *Journal of Intelligent Information Systems*, *43*(1), 183–203.
Mahmood, A., & Small, M. (2015). Subspace based network community detection using sparse linear coding. *IEEE Transactions on Knowledge and Data Engineering*, *28*(3), 801–812.
Mahmoud, H., Masulli, F., Rovetta, S., Russo, G. (2013). Community detection in protein-protein interaction networks using spectral and graph approaches. In *International meeting on computational intelligence methods for bioinformatics and biostatistics* (pp. 62–75). Cham: Springer.
McDaid, A.F., Greene, D., Hurley, N. (2013). Normalized mutual information to evaluate overlapping community finding algorithms arXiv:1110.2515v2.
Messaoudi, I., & Kamel, N. (2019). A multi-objective bat algorithm for community detection on dynamic social networks. *Applied Intelligence*, *49*(6), 2119–36.
Mitchell, M. (2006). Complex systems: network thinking. *Artificial Intelligence*, *170*(18), 1194–1212.
Newman, M.E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Science*, *98*(2), 404–409.
Newman, M.E. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577–8582.
Newman, M.E. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, *74*(3), 036104.
Newman, M.E. (2012). Communities, modules and large-scale structure in networks. *Nature Physics*, *8*(1), 25.
Ou, M., Cui, P., Pei, J., Zhang, Z., Zhu, W. (2016). Asymmetric transitivity preserving graph embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1105–111.4): ACM.
Palla, G., Derényi, I., Farkas, I., Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814–818.
Pattanayak, H.S., Sangal, A.L., Verma, H.K. (2019). Community detection in social networks based on fire propagation. *Swarm and evolutionary computation*, *44*, 31–48.
Ramezani, M., Khodadadi, A., Rabiee, H.R. (2018). Community detection using diffusion information. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *23;12*(2), 1–22.
Rashedi, E., Nezamabadi-pour, H., Saryazdi, S. (2009). GSA: A gravitational search algorithm. *Elsevier Information Sciences*, *179*, 2232–2248.
Rosvall, M., & Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, *105*(4), 1118–1123.
Sahebi, S., & Cohen, W.W. (2011). Community-based recommendations: a solution to the cold start problem. In *Workshop on recommender systems and the social web, RSWEB* (p. 60).
Saleh, M., Esa, Y., Mohamed (2019). A applications of complex network analysis in electric power systems. *Energies*, *11*(6), 1381.
Siddiquea, N., & Adelib, H. (2016). Applications of gravitational search algorithm in engineering.
Wang, D., Cui, P., Zhu, W. (2016). Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1225–1234): ACM.
Yadav, R.K., & Rai, A.K. (2020). Incorporating communities' structures in predictions of missing links. Journal of Intelligent Information Systems. https://doi.org/10.1007/s10844-020-00603-y.
Zachary, W.W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, *33*, 452–473.
Zhou, Z., & Amini, A. (2019). Analysis of spectral clustering algorithms for community detection: the general bipartite setting. *Journal of Machine Learning Research*, *20*(47), 1–47.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)
Galgotias University, Greater Noida, UP, India. Oct 2-4, 2020

Comparative Analysis of FPGA Based Hardware Design of Dynamic Time Warping Algorithm using Different Multiplier Architectures

Ritik Koul
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
ritik.koull@gmail.com

Mukul Yadav
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
ymukul1998@gmail.com

Kriti Suneja
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
kritisuneja@dtu.ac.in

***Abstract*—Dynamic Time Warping algorithm is used widely in speech recognition applications to provide a measure of similarity between two time-varying sequences. It finds its use in the speech processing applications due to the fact that the two sequences to be compared may be of different lengths and Euclidean distance computation may not give a true picture of the diversity between the two sequences. This algorithm involves a lot of computational complexity because of the large number of multiplication operations involved in its implementation. Also, the time delay is quite significant. In this work, hardware implementation of DTW matrix computation along with a comparative analysis using different multipliers for studying the hardware utilization statistics and time delay involved in the operation has been done. Verilog Hardware Description language (HDL) has been used and all the simulations have been performed on Model Sim.**

Keywords—Hardware Description Language (HDL), Dynamic Time Warping (DTW), Input-Output Blocks (IOB's), LUT (Look-up Tables), Carry Save Adder (CSA), Canonical Signed Digit (CSD).

I. INTRODUCTION

In simple words Dynamic time warping is an algorithm that is used to analyse the similarity between two time series of different lengths. These time series may be simple arrays but they may have different lengths. Now the question may arise that whether we really want to compare two time series of different lengths in any practical application. The answer is a big yes.

If we want to recognize the voice of an individual using a track of his sound and say, we have one scenario where we have the track of the individual's voice where he speaks "HEY" in some tone but in another scenario, he speaks the same word but in some other tone say "HEYYYYY". Then there comes the need of an algorithm where we can compare the two audio tracks of different lengths and be able to find and conclude that these have been spoken by the same individual.

Owing to this, Dynamic time warping algorithm has been used to compare two temporal sequences which may have different speeds or say different lengths and thus can be used in application such as voice recognition [1].

Also, DTW is very efficient in software computing. It has been used in Human Activity Recognition [2], Speech recognition system [3], Online Signature Verification [4], User independent hand gesture recognition [5] and finds many more applications.

Let us illustrate the idea of Dynamic time warping algorithm using an example. Suppose we have two different temporal data shown in red and blue colors. It is to be noted that these two sequences have different lengths.

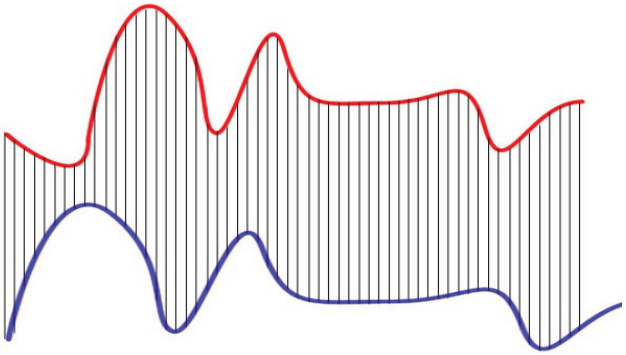


Fig. 1. Euclidean Distance Computation Approach [6]

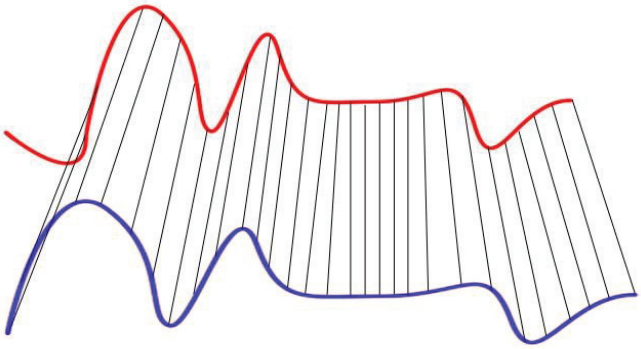


Fig. 2. Dynamic Time Warping Approach [6]

Although, we can see that the two waveforms follow similar pattern but the second one is longer than the first one. If we use one to one matching technique say Euclidean-distance computation for comparison of these two time series, we can see in the figure that some problem arises as this kind of matching is not perfectly synchronized and some part of the second waveform is left out and thus proper results will not be obtained.

Dynamic Time warping (DTW) algorithm overcomes this issue by avoiding one to one matching and instead uses one to many matching techniques. What it does, basically is

978-1-7281-5070-3/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:28:44 UTC from IEEE Xplore. Restrictions apply.

that, by using this one to many matching approaches, it aligns the crests and troughs properly so that these are matched perfectly and we have no portion of the curve left as it was in the previous case where one to one matching was being used.

II. METHODOLOGY

Dynamic Time Warping (DTW) algorithm finds the edit distance, which is the minimum number of editing operations required to convert one sequence into another, between two sequences by analysing three operations: substitution, insertion, and deletion [7]. Besides this, it is a search algorithm which finds an optimal warping path dynamically for the minimal cost alignment of input on template.

Measuring the edit distance is best reflected as a 2-dimensional grid with the two sequences on the top and left side respectively. The central idea is to formulate optimal path to any intermediate point X in the grid in terms of optimal paths of all its immediate antecedents. Let $\gamma(n,m)$ = Minimum path cost from origin to any point (n,m) in the grid.

$$\gamma(n,m) = d(n,m) + \min \{ \gamma(n-1,m), \gamma(n,m-1), \gamma(n-1,m-1) \} \quad (1)$$

where d(n,m) is the square of the Euclidean distance between the two elements of the target sequences at X . Hence, starting from the origin and computing minimum path cost for every grid point entry, proceeding from top left to bottom right corner provides the minimum edit distance as the last entry [8].

By using the above equation, we can implement a function in MATLAB to compute the minimum distance between the two sequence.

```
function [DTW,OP] = DTWCmp(A,B)

  n = length(A); m = length(B);

  DTW = zeros(n,m);
  DTW(1,1) = edsq(A(1),B(1));

  for j = 2 : m
     DTW(1,j) = edsq(A(1),B(j)) + DTW(1,j-1);
  end

  for j = 2 : n
     DTW(j,1) = edsq(A(j),B(1)) + DTW(j-1,1);
  end

  for i = 2 : n
     for j = 2 : m
        cost = edsq(A(i),B(j));
        DTW(i,j) = cost + min([ DTW(i-1,j),
                                DTW(i,j-1),DTW(i-1,j-1) ]);
     end
  end
```

The above part represents a section of MATLAB code that computes the minimum path cost to any point. Here A and B are the two sequences which are to be analyzed, edsq is a user defined function to calculate the square of the Euclidean distance. The first row and the first column are first calculated this require the previous element value and no comparisons. Using these values all the remaining values can be calculated as per the equation.

```
k = n; h = m;
OP = [DTW(k,h)];

while (k+h)~=2
   if k == 1
      value = DTW(k,h); h = h - 1;
   elseif h == 1
      value = DTW(k,h);  k = k - 1;
   else
      [value , number] = min([ DTW(k-1,h),
                               DTW(k,h-1),DTW(k-1,h-1)]);
      switch (number)
         case 1
            k = k - 1;
         case 2
            h = h - 1;
         case 3
            k = k - 1; h = h -1 ;
       end
    end
    OP = [OP,value];
 end
end
```

The above part represents a section of MATLAB code that computes the path starting from the very last element of the DTW matrix to the first element such that the sum of elements in the path is minimum. OP is a vector that contains the elements of the DTW matrix that forms the optimal warping path.

If we consider the two sequences as:

A = [1 1 2 2 3 3 3 4]; B = [8 7 4 3 2 1];

The DTW matrix for these two sequences can be computed using equation 1:

8	99	99	66	66	42	42	42	25
7	50	50	30	30	17	17	17	9
4	14	14	5	5	1	1	1	0
3	5	5	1	1	0	0	0	1
2	1	1	0	0	1	2	3	7
1	0	0	1	2	6	10	14	23
	1	1	2	2	3	3	3	4

Fig. 3. DTW Matrix for the two sequences

For computation of square of Euclidean distance between the two points, A[i] and B[i], we need to perform the operation:

$$E[i] = (A[i] - B[i]) * (A[i] - B[i]) \quad (2)$$

For the example considered above, we need to perform 48 multiplication operations for computation of DTW matrix. The number of multiplication operations required for this computation increases linearly as the length of the sequences increase.

Thus, for efficient hardware implementation of the algorithm, conventional array multiplication may not be a suitable technique.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:28:44 UTC from IEEE Xplore. Restrictions apply.

In this work we have integrated various multiplication techniques – Array Multiplier [9], Multiplier using Carry Save Addition technique, Wallace Multiplier [10], Dadda Multiplier [10], Vedic Multiplier [11,12] and Multiplier using Canonical signed digit representation of numbers[13], with the hardware implementation of the DTW matrix computation operation to find the most suitable implementation.

III. Simulation Results

Fig. 4 shows the results for the software implementation of the algorithm in MATLAB where Z represents the DTW matrix and X is a 1-D vector having elements that forms the part of the optimal warping path as explained above.

```
Z =

     0     0     1     2     6    10    14    23
     1     1     0     0     1     2     3     7
     5     5     1     1     0     0     0     1
    14    14     5     5     1     1     1     0
    50    50    30    30    17    17    17     9
    99    99    66    66    42    42    42    25

>> X

X =

    25     9     0     0     0     0     0     0     0     0
```

Fig. 4. MATLAB Results

Simulation results from Modelsim are as shown in Fig. 5.

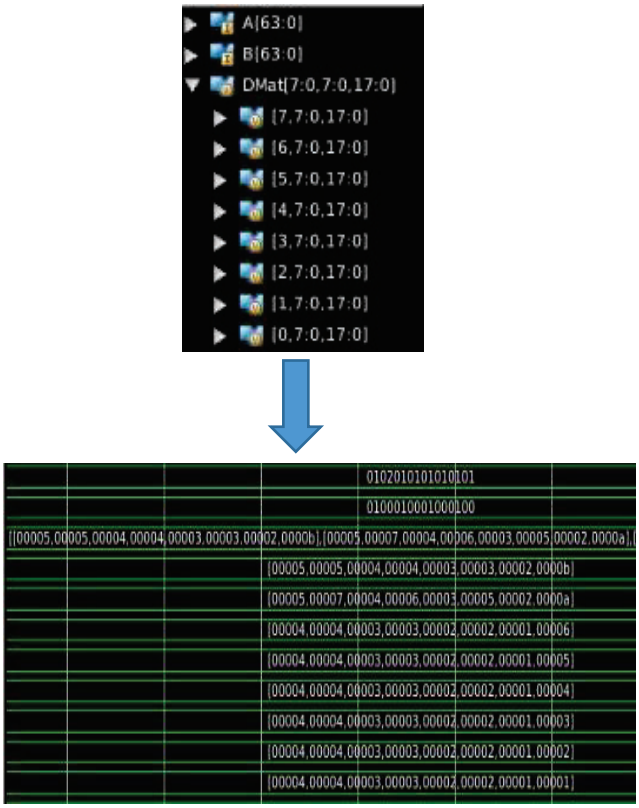


Fig. 5. Simulation Results in ModelSim

Fig. 6,7,8 and 9 show the hardware utilization statistics in terms of number of Slice LUT's being used for sequences of length 8,7,6 and 5 respectively.

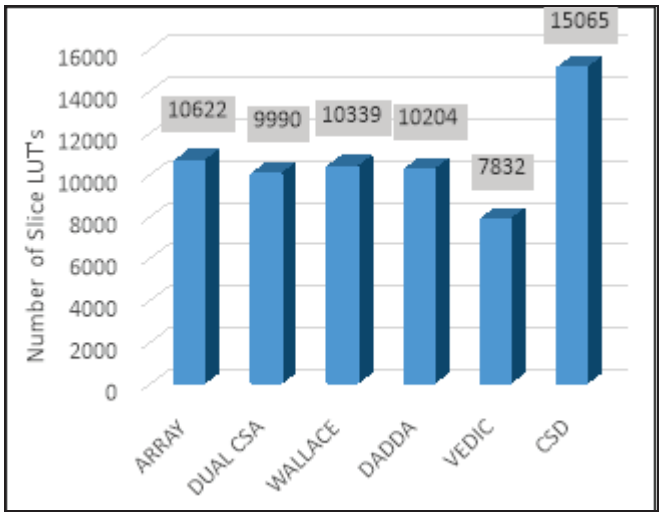


Fig. 6. No. of Slice LUT's used for sequences of length 8 for different multipliers

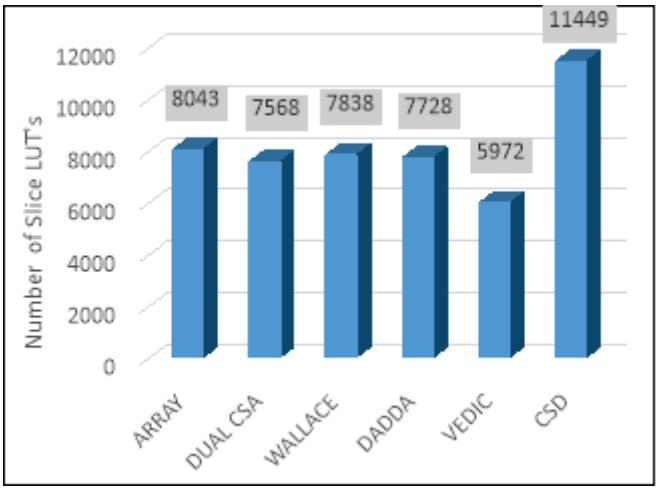


Fig. 7. No. of Slice LUT's used for sequences of length 7 for different multipliers

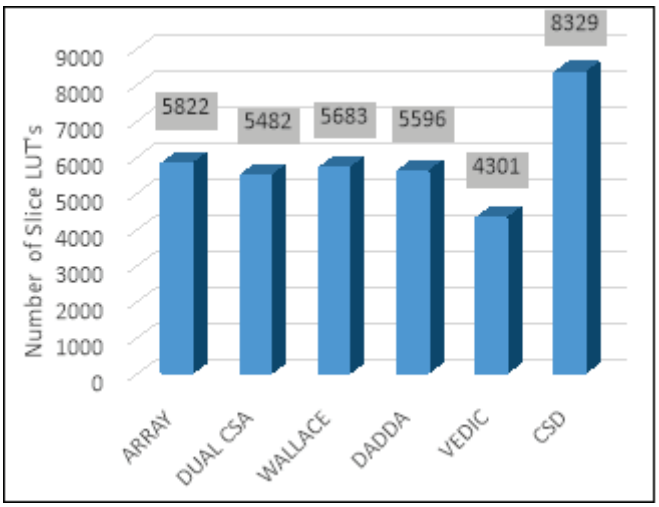


Fig. 8. No. of Slice LUT's used for sequences of length 6 for different multipliers

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:28:44 UTC from IEEE Xplore. Restrictions apply.

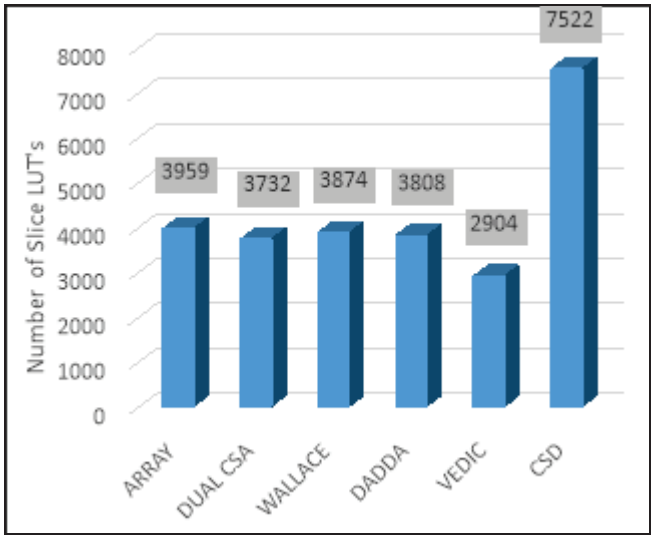


Fig. 9. No. of Slice LUT's used for sequences of length 5 for different multipliers

Figure 10,11,12 and 13 show the delay associated with sequences of length 8,7,6 and 5.

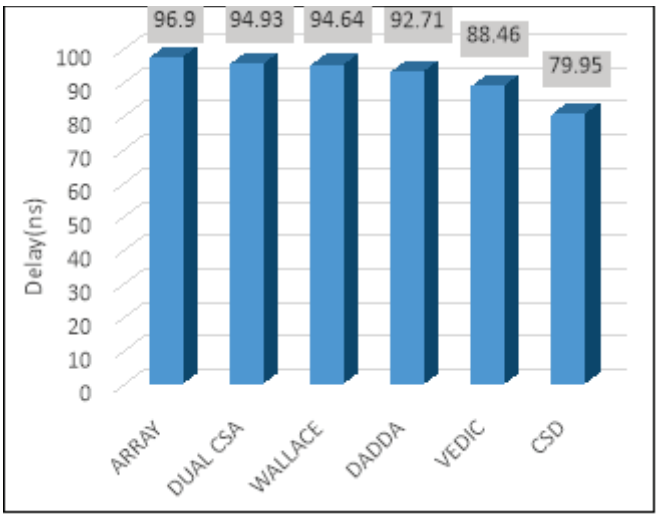


Fig. 10. Delay for sequences of length 8 for different multipliers

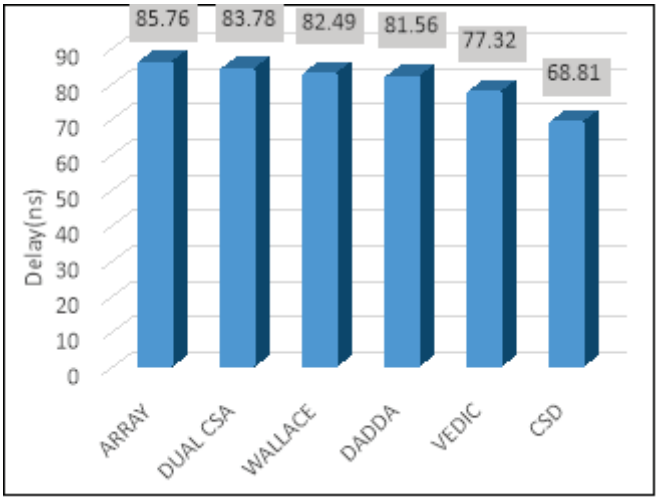


Fig. 11. Delay for sequences of length 7 for different multipliers

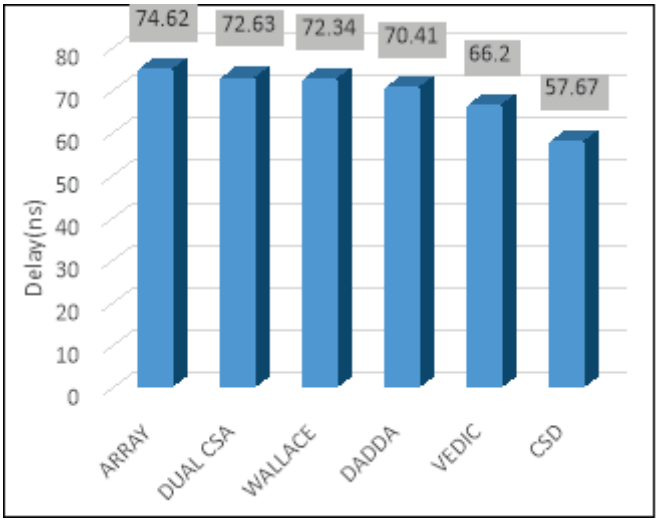


Fig. 12. Delay for sequences of length 6 for different multipliers

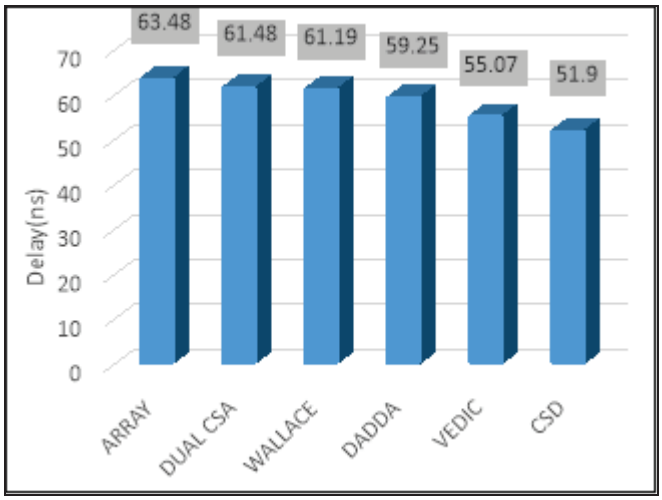


Fig. 13. Delay for sequences of length 5 for different multipliers

Table 1 summarizes the variation in the hardware utilization statistics derived upon varying the length of the two sequences being considered and the integrated multiplication technique.

TABLE I. VARIATION OF SLICE LUT'S

SEQUENCE LENGTH	ARRAY	DUAL CSA	WALLACE	DADDA	VEDIC	CSD
8 x 8	10622	9990	10339	10204	7832	15065
7 x 7	8043	7568	7838	7728	5972	11449
6 x 6	5822	5482	5683	5596	4301	8329
5 x 5	3959	3732	3874	3808	2904	7522

Table 2 summarizes the variation in the delay involved in the operation derived upon varying the length of the two sequences being considered and the integrated multiplication technique. All of the delays have been measured in nano-seconds(ns).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:28:44 UTC from IEEE Xplore. Restrictions apply.

TABLE II. VARIATION OF DELAY

SEQUENCE LENGTH	ARRAY	DUAL CSA	WALLACE	DADDA	VEDIC	CSD
8 x 8	96.905	94.937	94.646	92.713	88.467	79.95
7 x 7	85.766	83.786	82.495	81.561	77.328	68.81
6 x 6	74.626	72.634	72.343	70.41	66.202	57.671
5 x 5	63.486	61.482	61.191	59.258	55.075	51.9

The trend for the number of bonded IOB's being utilized upon variation of the sequence length is as shown in Fig. 14.

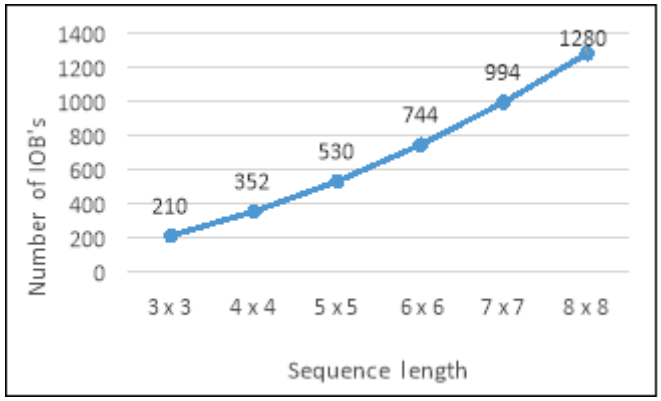


Fig. 14. Variation of IOB's

It can be clearly deduced from the above graphs and tables that the incorporation of conventional Array multiplication techniques into the design of DTW Algorithm leads to a lot of inefficiency in terms of delay involved in the operation.

Use of other multiplication techniques such as multipliers involving carry save addition technique to sum the partial products, Wallace Multipliers and Dadda multipliers, which make use of compressors to achieve the result of the multiplication process in less time, we can improve the design by incorporating them in the design.

Also, use of such multipliers gives the additional benefit of enhancing hardware efficiency in terms of number of the Slice LUT's being used for the synthesis of the design.

Another multiplication technique involving Vedic mathematics gives significant advantage, both in terms of delay and hardware utilization statistics.

Most of the designs developed revolve around these 5 techniques. However as shown in this work, multiplier that involves representation of numbers using Canonical Signed Digit representation, if incorporated in the design, can considerably reduce the delay involved in the operation. It leads all the other multiplication techniques in this respect. However, there is some trade-off involved in the design.

The design that incorporates Canonical Signed Digit multiplier makes use of a greater number of Slice LUT's. Thus, in terms of hardware utilization, it may appear that this technique is not too good. However, in modern designs, delay is the most significant parameter that needs to be minimized and most of the modern FPGA's do consist a large number of Slice LUT's, so CSD multiplication technique can be integrated to develop efficient Dynamic Time Warping algorithm as an independent module or a co-processor.

IV. CONCLUSION AND FUTURE WORK

This work mainly focusses on the implementation of Dynamic Time Warping algorithm. It is seen that the algorithm requires a large number of multipliers proportional to the t length of the sequences. The delay involved in the multiplication operation is quite significant and thus there is a need to integrate the DTW algorithm with multipliers which are efficient in terms of delay involved. By integrating the DTW algorithm with the Canonical Signed Digit multiplier, it is deduced that we can reduce the delay involved significantly. Further this system may be integrated with other larger systems which incorporate DTW algorithm as a part of their functionality to develop faster systems.

REFERENCES

[1] H.Mansour, Abdelmajid & Zen Alabdeen Salh, Gafar & Mohammed, Khalid. (2015). Voice Recognition using Dynamic Time Warping and Mel-Frequency Cepstral Coefficients Algorithms. International Journal of Computer Applications. 116. 34-41. 10.5120/20312-2362.

[2] M. Masnad, G. M. MukitHasan, K. M. Iftekhar and M. S. Rahman, "Human Activity Recognition using DTW Algorithm," 2019 IEEE Region 10 Symposium (TENSYMP), Kolkata, India, 2019, pp. 39-43.

[3] Zhang Jing and Zhang Min, "Speech recognition system based improved DTW algorithm," 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering, Changchun, 2010, pp. 320-323

[4] B. V. Patil and P. R. Patil, "An Efficient DTW Algorithm for Online Signature Verification," 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), Sangamner, 2018, pp. 1-5

[5] Shah Muhammed Abid Hussain and A. B. M. Harun-ur Rashid, "User independent hand gesture recognition by accelerated DTW," 2012 International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, 2012, pp. 1033-1037

[6] https://sflscientific.com/data-science-blog/2016/6/3/dynamic-time-warping-time-series-analysis-ii

[7] M. Ravishankar, "Isolated word recognition using dynamic time warping," Design and implementation of ASR systems. http://asr.cs.cmu.edu/spring2010/class3/class3.25jan.dtw.pdf

[8] Kriti Suneja, Dr. Malti Bansal , " Hardware Design of Dynamic Time Warping Algorithm based on FPGA in Verilog," International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)Volume 4, Issue 2, February 2015.

[9] Kumar, Aniket & Vishikha,. (2017). COMPARATIVE ANALYSIS OF VEDIC & ARRAY MULTIPLIER. 8. 17-27

[10] An Approximate Multiplier for Efficient Multiplication using Xilinx http://www.ijitee.org/wp-content/uploads/papers/v9i3/C8128019320

[11] Premananda B.S.,Samarth S. Pai, Shashank B, Shashank S. Bhat," Design and Implementation of 8-Bit Vedic Multiplier," IJAREEIE, Volume. 2, Issue 12, December 2013

[12] S. Garg, S. Garg and V. Sachdeva, "Comparative analysis of 8 X 8 Bit Vedic and Booth Multiplier," 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), New Delhi, 2014, pp. 2607-2610.

[13] Koul, Ritik & Yadav, Mukul & Suneja, Kriti. (2020). Novel FPGA-Based Hardware Design of Canonical Signed Digit Matrix Multiplier and Its Comparative Analysis with Other Multipliers. 10.1007/978-981-15-1059-5_8.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:28:44 UTC from IEEE Xplore. Restrictions apply.

Control and Operation of a Small Autonomous DC Microgrid with Improved Dynamic Performance

Utsav Patel
School of Technology
Pandit Deendayal Petroleum University
Gandhinagar, India
utsavrpatel29@gmail.com

Avirup Maulik
Department of Electrical Engineering
Delhi Technological University
Delhi, India
avirupmaulik@gmail.com

Debapriya Das
Department of Electrical Engineering
Indian Institute of Technology
Kharagpur, India
ddas@ee.iitkgp.ac.in

***Abstract*—An islanded DC Microgrid comprises dispatchable and renewable generation units, battery energy storage systems, and loads. An islanded DC microgrid presents several operational challenges due to load and renewable generation fluctuations. Therefore, a dynamic performance analysis is required to ensure that a DC Microgrid operates satisfactorily. In this study, we have analyzed the dynamic performance of an islanded DC microgrid in detail. Based on the dynamic analysis, we have proposed a primary control method for improving the system dynamic response. We have derived a small-signal model of the entire DC Microgrid and carried out an eigenvalue and participation factor analysis. Based on the participation factor analysis, suitable system states are identified, and a controller is designed for a better transient performance of the system. Time-domain simulations in EMTDC/PSCAD are carried out to verify and establish that the method proposed in this paper yields a satisfactory system performance.**

***Index Terms*—DC Microgrid, Small-signal model, eigenvalue, participation factor, dynamic performance**

I. Introduction

A DCMG comprises both renewable and dispatchable power generating sources, storage units, and loads, and can operate either in conjunction with the main grid (grid-connected operation) or in the standalone mode (islanded/autonomous operation) [1]. The renewable generating units are controlled using the "maximum power point tracking" (MPPT) algorithm, while a hierarchical control strategy was proposed for controlling dispatchable generating units [2]. Since most of the generating units, as well as storage devices in a DCMG, is converter interfaced, it suffers from a lack of system inertia. The lack of system inertia presents a severe operational challenge to the system due to rapid fluctuation of the system voltage to disturbances like load and renewable generation variation.

Some research related to control strategies of a DCMG for satisfactory system performance has been reported in the literature. Several virtual inertia based schemes were proposed to improve the dynamic performance of a DCMG [3]–[10]. We observe that only a limited amount of research work has been done in this area. Moreover, in all of the referred papers [3]–[10], researchers focused on implementing virtual inertia based schemes. Alternative control schemes to improve the performance of a DCMG has not been thoroughly explored.

In the present paper, we have extended the above threads of work. We have proposed a control technique for the primary control level, which is an alternative to the reported virtual-inertia based techniques, to improve the system performance under dynamic conditions. We have formulated the complete "small-signal model" of the entire system, and carried out eigenvalue and participation factor analysis to identify suitable states for controlling to improve the dynamic performance. Time-domain simulations on EMTDC/PSCAD platform are carried out to implement and check that the control approach proposed in this work improves the performance of the system under dynamic conditions. The arrangement of the remaining portions of this paper is as follows. In section II, the system model including the problem formulation is discussed along with the small-signal models of individual components, and the overall system small-signal model. In section III, we have discussed simulation results, followed by conclusions in section IV.

II. System model and problem formulation

The autonomous DCMG system considered in this paper comprises a photovoltaic (PV) generating source, a battery energy storage system (BESS), and a dispatchable generator feeding a load connected to the common DC bus. There are two buses in the system. At bus 2, the PV unit and the BESS are connected, along with a load. At bus 1, the dispatchable unit is connected along with a load. A feeder connects the two buses. The system model is shown in Fig.1. The PV unit is interfaced with the DCMG through a boost converter. A "perturb and observe" algorithm is used in the MPPT controller. A Lithium-Ion BESS is interfaced through a bi-directional buck-boost converter. The dispatchable unit is interfaced via a buck converter. The dc bus voltage of test system is 60 volts. Battery nominal voltage is 48 volts and rated capacity is 8 Ah. PV rating is 500 Watt.

A. Small-signal model

All variables are time dependent. All hatted quantities mean small departures about steady-state quiescent points.

1) Boost Converter: The equations for the small-signal model of the boost converter (shown in Fig.2) are as follows [11]:

$$\frac{d\hat{i}_l}{dt} = -\frac{R_l}{L}\hat{i}_l - \frac{D'}{L}\hat{v}_0 + \frac{1}{L}\hat{v}_s + \frac{V}{L}\hat{d} \tag{1}$$

978-1-7281-6350-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

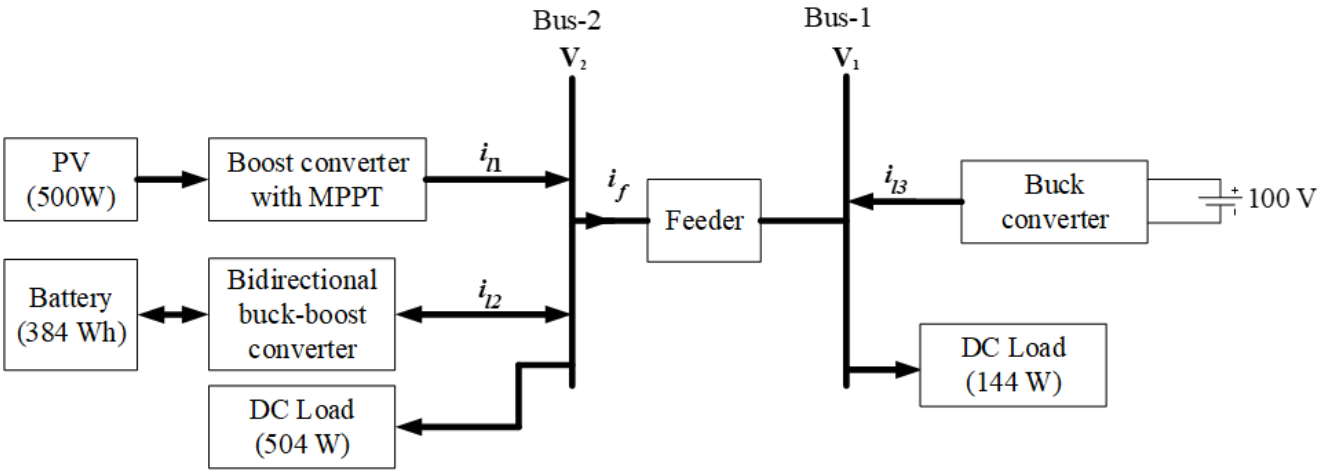


Fig. 1: DCMG system model

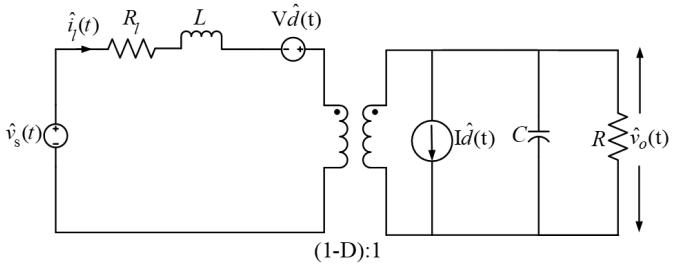


Fig. 2: Boost converter - Small-signal model with averaged dynamics [11]

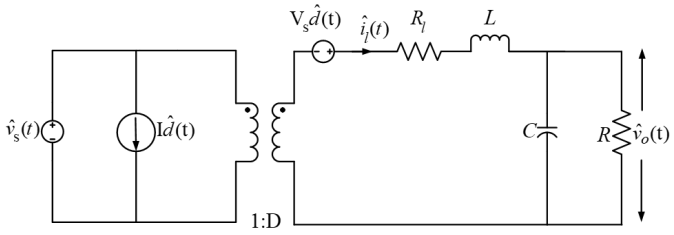


Fig. 3: Buck converter - Small-signal model with averaged dynamics [11]

$$\frac{d\hat{v}_0}{dt} = \frac{D'}{C}\hat{i}_l - \frac{1}{RC}\hat{v}_0 - \frac{I}{C}\hat{d} \tag{2}$$

where, i_l is the inductor current, R_l is the switching and the stray resistance in the inductor. L and C denote the inductance and capacitance, respectively. R is the load resistance, d is the duty ratio, D is the quiescent value of the duty ratio, $D' = 1 - D$, v_0 is the output voltage, V is the quiescent value of output voltage, I is the quiescent value of the inductor current.

2) Buck converter: The equations are as follows [11] (please see Fig.3):

$$\frac{d\hat{i}_l}{dt} = -\frac{R_l}{L}\hat{i}_l - \frac{1}{L}\hat{v}_0 + \frac{D}{L}\hat{v}_s + \frac{V_s}{L}\hat{d} \tag{3}$$

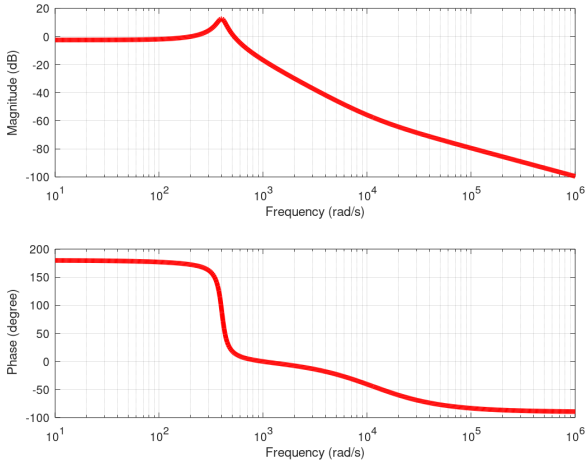


Fig. 4: Bode plot of the bi-directional converter before compensation

$$\frac{d\hat{v}_0}{dt} = \frac{1}{C}\hat{i}_l - \frac{1}{RC}\hat{v}_0 \tag{4}$$

3) Feeders: The equation governing the small-signal model of a feeder is as follows:

$$\hat{v}_i - \hat{v}_j = L_f\frac{d\hat{i}_f}{dt} + R_f\hat{i}_f \tag{5}$$

where v_i is the voltage of the i^{th} bus, L_f and R_f denote the feeder inductance and resistance respectively. i_f denotes the feeder current.

4) Controller structure: The primary control loop with the proposed control strategy is implemented on both the converters (i.e., the bi-directional buck-boost dc/dc converter interfacing the BESS unit, and the buck converter interfacing the dispatchable unit). The modified primary control loop is shown in Fig.6. The current of the converter ($\hat{i}_L$) is fed back through a derivative controller for the battery converter to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

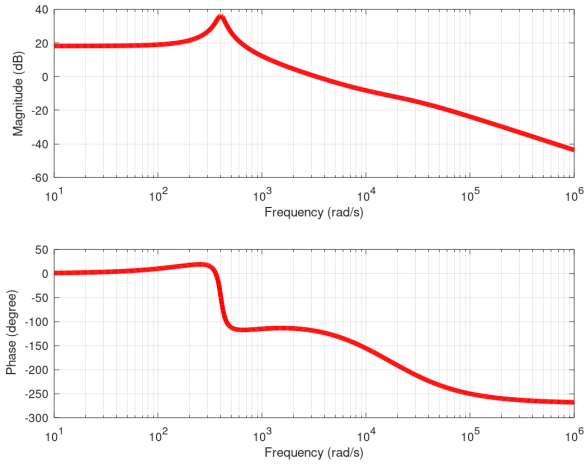


Fig. 5: Bode plot of the bi-directional converter after compensation

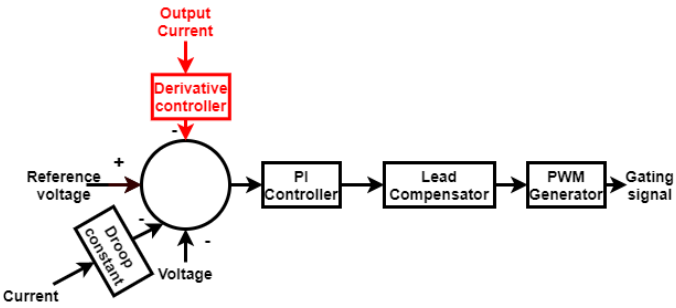


Fig. 6: Proposed control scheme at the primary level for battery charging and discharging

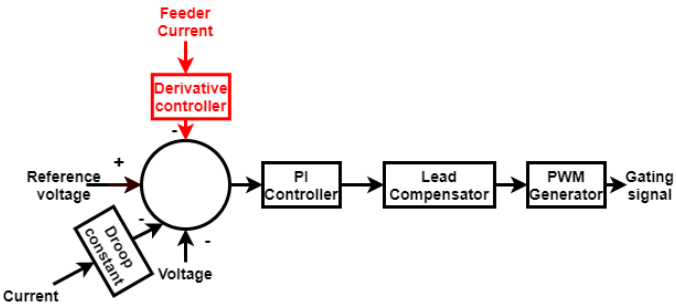


Fig. 7: Proposed control scheme at the primary level for the dispatchable unit

improve the dynamic performance of the system. Also, droop control is implemented by feeding a voltage proportional to the converter current. The control structure comprises a PI controller to reduce the steady-state voltage error to zero. The output of the PI controller is fed to a lead compensator. The lead compensator is provided and designed to achieve a phase-margin of $55°$ for the converter, and a cross-over frequency around $500Hz$ (i.e., $\frac{1}{20}$ times the switching frequency of 10 kHz). The phase-margin of the converter before compensation is $-15°$, and the cross-over takes place at 83.41 Hz.The bode plot before compensation is shown in Fig.4. With compensation, the phase-margin increases to $60°$, and the cross-over frequency is 526.94 Hz. The bode plot after compensation is shown in Fig. 5. The PI controller is designed to have a cut-off frequency of 50 Hz (i.e., $\frac{1}{10}^{th}$ of the cross-over frequency of the lead compensator). The controller for the buck converter also has the same topology and design philosophy, but the modification involves addition of a negative derivative feedback of the feeder current. Modified primary control loops are shown in Fig.6, and Fig.7 respectively. Modifications in the primary control loops were carried out based on participation factor analysis.

5) Complete system model: Individual system models are combined to obtain the "small-signal model" of the complete system (Fig.8). The states are marked on the figure. Note that $\hat{v}_1 = \hat{x}_1$; $\hat{v}_2 = \hat{x}_5$;$\hat{i}_1 = \hat{x}_2$; $\hat{i}_2 = \hat{x}_6$; $\hat{i}_f = \hat{x}_9$. Since the PV unit operates on MPPT mode, the change of PV current injected into the system is considered as a disturbance. The following equations govern the small-signal model of the overall system:

$$\dot{\hat{x}}_1 = -\frac{1}{R_1C_1}\hat{x}_1 + \frac{1}{C_1}\hat{x}_2 + \frac{1}{C_1}\hat{x}_9 \tag{6}$$

$$\dot{\hat{x}}_2 = -\frac{1}{L_1}\hat{x}_1 + \frac{V_{g1}}{V_M L_1}\hat{x}_3 + \frac{D_1}{L_1}\hat{v}_{g1} \tag{7}$$

$$\begin{aligned}\dot{\hat{x}}_3 = &\sum_{ii\in S_1, jj\in S_2} \frac{k_6}{k_8}T_{ii}\hat{x}_{jj} - \frac{1}{k_8}\hat{x}_3 + \frac{k_7}{k_8}\hat{x}_4 \\ &- \frac{k_6K_9k_{d1}}{k_8L_fC_2k_{10}}\hat{i}_{PV} + \frac{k_6}{k_8k_{10}}\hat{v}_1^{ref} \\ &+ \frac{k_7k_9Rd_1D_1}{k_8k_{10}L_1}\hat{v}_{g1}\end{aligned} \tag{8}$$

where $S_1 = [7, 8, 0, 9, 10, 11, 12]$; $S_2 = [1, 2, 3, 5, 6, 7, 9]$.

$$\begin{aligned}\dot{\hat{x}}_4 = &\sum_{ii\in S_1, jj\in S_2} T_{ii}\hat{x}_{jj} - \frac{1}{k_8}\hat{x}_3 + \frac{k_7}{k_8}\hat{x}_4 \\ &- \frac{k_9k_{d1}}{L_fC_2k_{10}}\hat{i}_{PV} + \frac{1}{k_{10}}\hat{v}_1^{ref} + \frac{k_9Rd_1D_1}{L_1k_{10}}\hat{v}_{g1}\end{aligned} \tag{9}$$

$$\dot{\hat{x}}_5 = \frac{-1}{R_2C_2}\hat{x}_5 + \frac{D_2{}'}{C_2}\hat{x}_6 - \frac{I_2}{C_2V_M}\hat{x}_7 - \frac{1}{C_2}\hat{x}_9 + \frac{1}{C_2}\hat{i}_{PV} \tag{10}$$

$$\dot{\hat{x}}_6 = \frac{-D_2{}'}{L_2}\hat{x}_5 + \frac{V_2}{L_2V_M}\hat{x}_7 + \frac{1}{L_2}\hat{v}_{g2} \tag{11}$$

$$\begin{aligned}\dot{\hat{x}}_7 = &\sum_{ii\in S_3, jj\in S_4} \frac{k_1}{k_3}T_{ii}\hat{x}_{jj} + \frac{k_1T_{20}-1}{k_3}\hat{x}_7 + \frac{k_1T_{21}+k_2}{k_3}\hat{x}_8 + \\ &\frac{k_1}{k_3(k_5+k_{d2}k_4T_{17})}\hat{v}_2^{ref} + \frac{k_1T_{24}}{k_3}\hat{v}_{g2} + \frac{k_1T_{25}}{k_3}\hat{i}_{PV}\end{aligned} \tag{12}$$

where $S_3 = [18, 19, 22]$, $S_4 = [5, 6, 9]$

$$\begin{aligned}\dot{\hat{x}}_8 = &\sum_{ii\in S_5, jj\in S_6} T_{ii}\hat{x}_{jj} + \frac{1}{k_5+k_{d2}k_4T_{17}}\hat{v}_2^{ref} - \\ &T_{24}\hat{v}_{g2} + T_{25}\hat{i}_{PV}\end{aligned} \tag{13}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

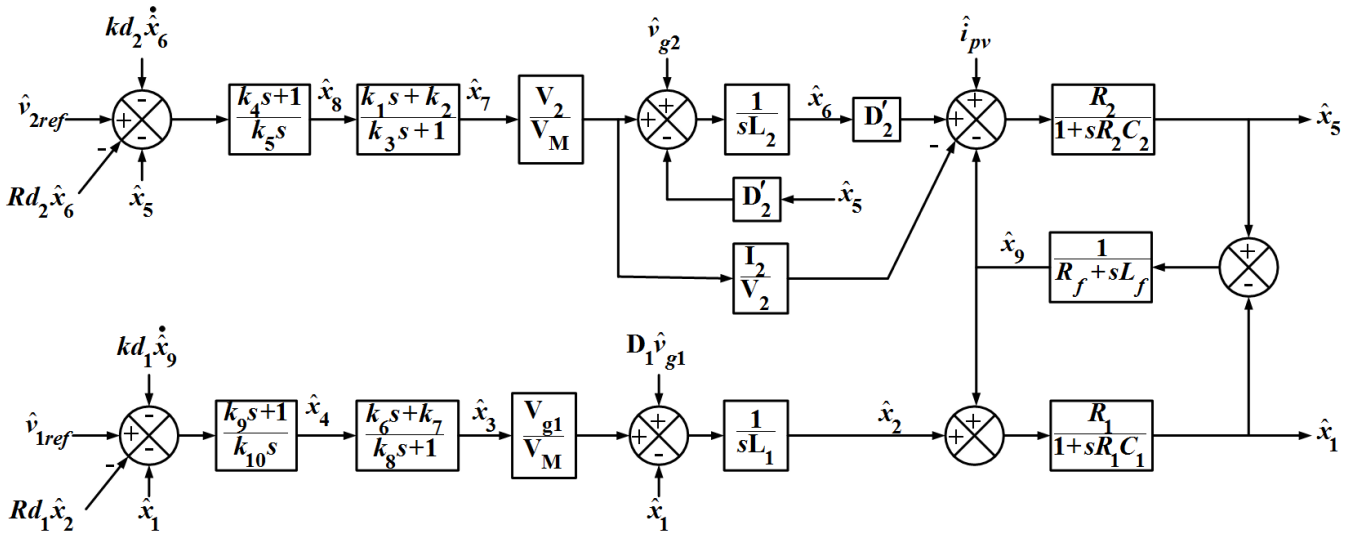


Fig. 8: Small-signal model of the overall system

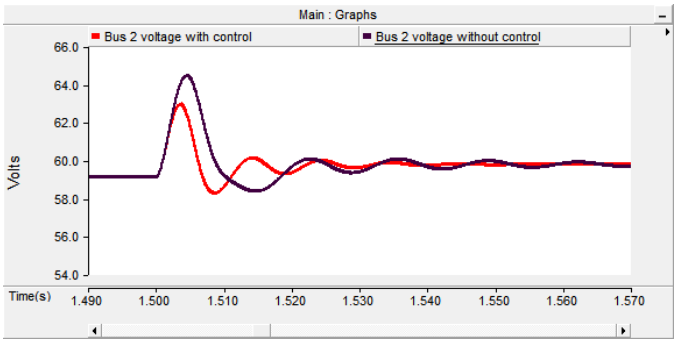


Fig. 9: Bus 2 voltage with and without the proposed controller in the vicinity of $t = 1.5s$

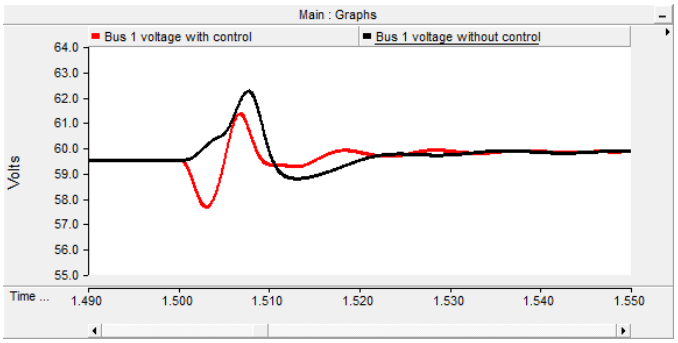


Fig. 10: Bus 1 voltage with and without the proposed controller in the vicinity of $t = 1.5s$

where $S_5 = [18, 19, 20, 21, 22]$, $S_6 = [5, 6, 7, 8, 9]$

$$\dot{\hat{x}}_9 = \frac{-1}{L_f}\hat{x}_1 + \frac{1}{L_f}\hat{x}_5 - \frac{R_f}{L_f}\hat{x}_9 \tag{14}$$

The state-space equations of the system before the modification in the control loops can be found out by substituting $k_{d1} = k_{d2} = 0$.

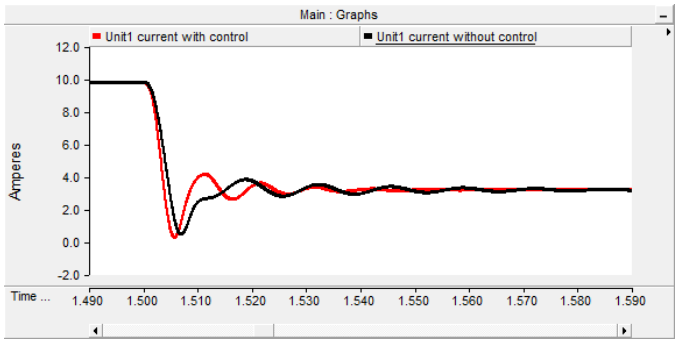


Fig. 11: Converter 1 current with and without the proposed controller in the vicinity of $t = 1.5s$

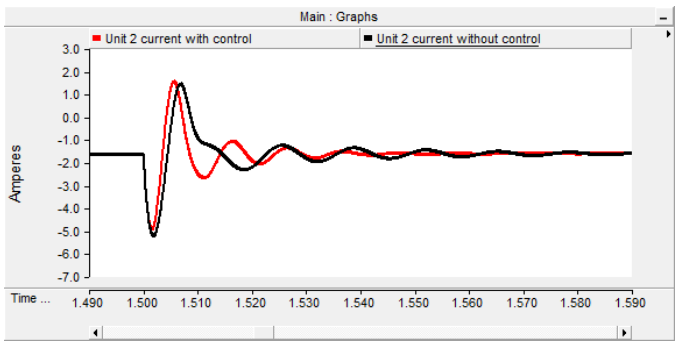


Fig. 12: Converter 2 current with and without the proposed controller in the vicinity of $t = 1.5s$

6) Practical implementation of the derivative controller: The practical implementation of a derivative controller is difficult since it amplifies high-frequency noises. An actual realization employs passing the feedback signal through a low pass filter (LPF), followed by a washout filter. A washout filter is a high pass filter and provides zero static gain. More on the design aspect of a washout filter can be found in [12].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Eigenvalue and participation factors before implementing the modified control strategy

States	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
	-23602.02	-16883.01	$-1547.13 \pm j1128.73$	$-67.08 \pm j640.08$	-174.75	$-16.06 \pm j23.68$
$\hat{x}_1$	0.000	0.003	0.225	0.041	0.393	0.017
$\hat{x}_2$	0.000	0.158	0.469	0.062	0.066	0.015
$\hat{x}_3$	0.000	0.837	0.054	0.002	0.003	0.000
$\hat{x}_4$	0.000	0.002	0.168	0.032	0.413	0.012
$\hat{x}_5$	0.000	0.000	0.006	0.438	0.016	0.023
$\hat{x}_6$	0.001	0.000	0.000	0.201	0.062	0.318
$\hat{x}_7$	0.998	0.000	0.000	0.001	0.000	0.001
$\hat{x}_8$	0.000	0.000	0.000	0.016	0.005	0.434
$\hat{x}_9$	0.000	0.000	0.077	0.207	0.043	0.180

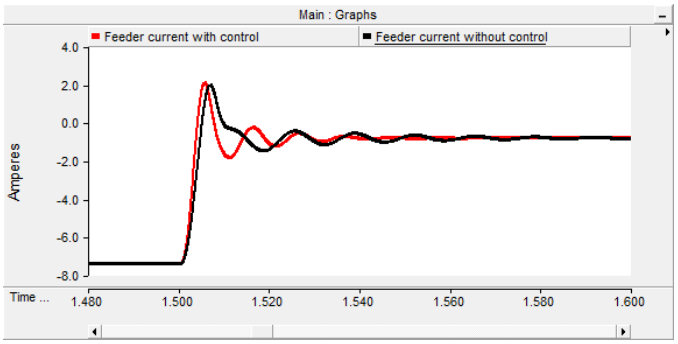


Fig. 13: Feeder current with and without the proposed controller in the vicinity of $t = 1.5s$

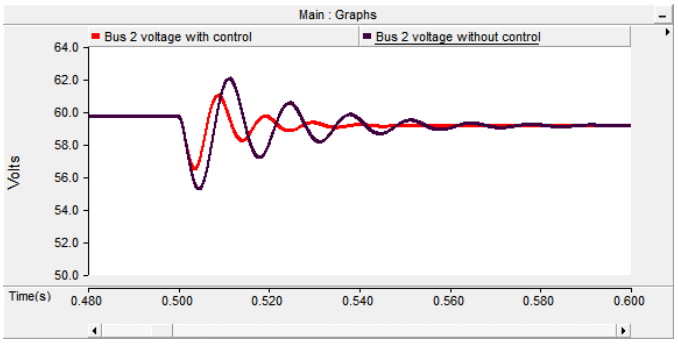


Fig. 14: Bus 2 voltage with and without the proposed controller in the vicinity of $t = 0.5s$

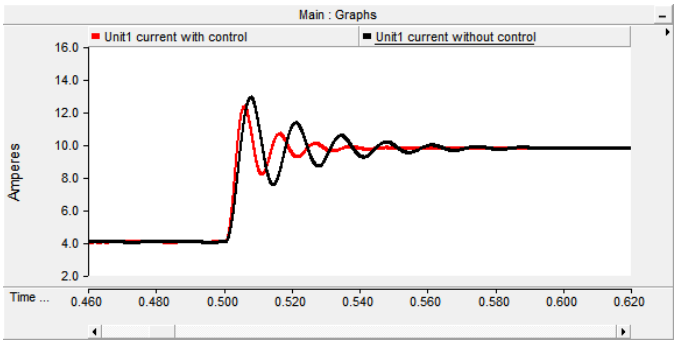


Fig. 15: Converter 1 current with and without the proposed controller in the vicinity of $t = 0.5s$

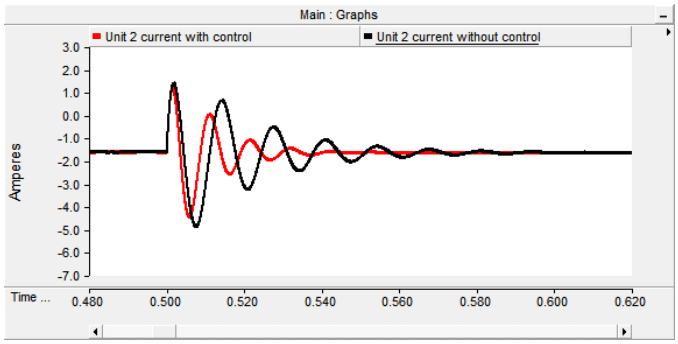


Fig. 16: Converter 2 current with and without the proposed controller in the vicinity of $t = 0.5s$

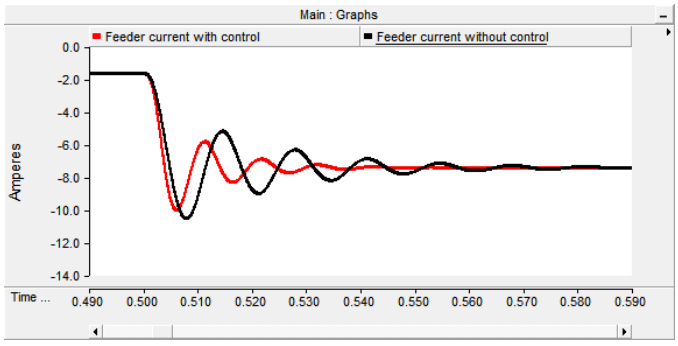


Fig. 17: Feeder current with and without the proposed controller in the vicinity of $t = 0.5s$

B. *Objective/Problem formulation*

The dynamic system performance can be improved from the small-signal model after finding the eigenvalues and by choosing controller parameters so as to shift the most dominant eigenvalue to the left half of the s-plane and improving the system damping by the following [10]:

$$Minimize(max \underset{k \in NS}{(} Real(\lambda_k)) + \sum_{k \in NS} (1 - \xi_k)) \quad (15)$$

where λ_k and ξ_k are the k^{th} eigenvalue and the damping ratio respectively, NS is the number of system states.

III. Simulation Results

A. *Eigenvalue and participation factor analysis*

The system state-space equations before implementing the feedback control is obtained by substituting $k_{d1} = 0; k_{d2} = 0$.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

It is necessary to identify the states which participate most heavily in the oscillating modes. States which participate most heavily in the oscillating modes are selected for feedback control. The eigenvalues and participation of different states are given in Table I. It is observed that modes 3,4, and 6 are oscillatory. Damping ratios of modes 3,4 and 6 are 0.808, 0.104, and 0.561, respectively. Damping of modes 4 and 6 are the least. Modes 4 and 6 should be damped to improve the dynamic response of the system. We further observe that states $\hat{x}_9$, and $\hat{x}_6$ participate heavily in the oscillatory modes 4, and 6. Therefore, states $\hat{x}_9$ (feeder current) and $\hat{x}_6$ (inductor/output current of the bi-directional converter) are chosen for controlling. Derivatives of states $\hat{x}_6$ and $\hat{x}_9$ are fed as negative feedback to the converters for improving the dynamic response of the system. It is found that $k_{d1} = 0.001$, and $k_{d2} = 0.5$ give a satisfactory performance.

B. Time domain simulations

Time-domain simulations in EMTDC/PSCAD are carried out to assess and validate the performance of the proposed control scheme. The sequence of events is as follows:

- For $0 \leq t < 0.5 \quad s$, the load at bus 1 is 144 W, the load at bus 2 is 544 W, and the solar PV unit generates a power of 500 W.
- At $t = 0.5 \quad s$, the solar generation drops to 150 W.
- At $t = 1.5 \quad s$, the solar generation becomes 300 W, and the load at bus 2 becomes 252 W.
- At $t = 2.5 \quad s$, the load at bus 1 becomes 72 W.

Since the disturbance at $t = 1.5s$ is the most severe (two simultaneous changes occurring in the system), we present the plots of system voltages and currents in the vicinity of this time instant. Plots of different system quantities without and with the controller are shown if Figures 9, 10, 11, 12, and 13. Further, system responses in the vicinity of $t = 0.5 \quad s$ (i.e., when the power generation from the PV unit at bus 2 changes) are also shown in Fig.14, Fig.15, Fig.16, and Fig.17. In all the cases, we observe that the oscillations in the system variables are damped out effectively using the proposed control strategy. The proposed control strategy is easy to implement on a hardware setup since it does not involve any complicated or burdensome computational algorithms.

IV. CONCLUSIONS

In this paper, we have carried out small-signal modeling of a small autonomous DCMG comprising a dispatchable unit, a BESS unit, and a PV source. From the eigenvalue and participation factor analysis, we have identified suitable system states for controlling so that the dynamic performance of the system is improved. From participation factor analysis, the feeder current and the output/inductor current of the bi-directional converter are identified as suitable states for controlling. An approximation of derivative controllers using washout filters is implemented. Time-domain simulations carried out on the EMTDC/PSCAD platform validates the effectiveness of the proposed control strategy in improving the system dynamic performance.

REFERENCES

[1] A. Maulik and D. Das, "Application of linearised load flow method for droop-controlled dcmgs," *IET Generation, Transmission & Distribution*, vol. 14, no. 6, pp. 1114–1126, 2020.

[2] Q. Shafiee, T. Dragičević, J. C. Vasquez, and J. M. Guerrero, "Hierarchical control for multiple dc-microgrids clusters," *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 922–933, 2014.

[3] E. Unamuno and J. A. Barrena, "Design and small-signal stability analysis of a virtual-capacitor control for dc microgrids," in *2017 19th European Conference on Power Electronics and Applications (EPE'17 ECCE Europe)*. IEEE, 2017, pp. P–1.

[4] C. Wang, J. Meng, Y. Wang, and H. Wang, "Adaptive virtual inertia control for dc microgrid with variable droop coefficient," in *2017 20th International Conference on Electrical Machines and Systems (ICEMS)*. IEEE, 2017, pp. 1–5.

[5] S. Samanta, J. P. Mishra, and B. K. Roy, "Implementation of a virtual inertia control for inertia enhancement of a dc microgrid under both grid connected and isolated operation," *Computers & Electrical Engineering*, vol. 76, pp. 283–298, 2019.

[6] X. Huang, Y. Hong, J. Liu, and P. Jia, "Virtual inertia online control strategy for dc grid based on variable droop coefficient," 2019.

[7] S. Samanta, J. P. Mishra, and B. K. Roy, "Virtual dc machine: an inertia emulation and control technique for a bidirectional dc–dc converter in a dc microgrid," *IET Electric Power Applications*, vol. 12, no. 6, pp. 874–884, 2018.

[8] X. Zhu, F. Meng, Z. Xie, and Y. Yue, "An inertia and damping control method of dc-dc converter in dc microgrid," *IEEE Transactions on Energy Conversion*, 2019.

[9] S. Samanta, J. P. Mishra, and B. K. Roy, "Ac load bus frequency control of a dc microgrid based on dc voltage regulation using inertia emulation and economic power management," *IET Generation, Transmission & Distribution*, vol. 13, no. 22, pp. 5117–5128, 2019.

[10] A. Maulik, D. Das, and A. Jain, "Improvement of the dynamic performance of an islanded dc microgrid using optimized virtual inertia," in *2019 21st International Middle East Power Systems Conference (MEPCON)*. IEEE, 2019, pp. 244–249.

[11] R. W. Erickson and D. Maksimovic, *Fundamentals of power electronics*. Springer Science & Business Media, 2007.

[12] Y. Wang, C. Wang, L. Xu, J. Meng, and Y. Hei, "Adjustable inertial response from the converter with adaptive droop control in dc grids," *IEEE Transactions on Smart Grid*, 2018.

APPENDIX

OVERALL SMALL-SIGNAL MODEL: VARIABLE EXPANSION

$T_0 = \frac{-k_9 Rd_1 V_{g1}}{V_M L_1}$; $T_1 = \frac{1}{L_f R_1 C_1} + \frac{R_f}{L_f^2}$; $T_2 = \frac{-1}{L_f C_1}$; $T_3 = \frac{-1}{L_f R_2 C_2} - \frac{R_f}{L_f^2}$; $T_4 = \frac{D_2{}'}{L_f C_2}$; $T_5 = \frac{-I_2}{L_f C_2 V_M}$; $T_6 = \frac{-1}{L_f C_1} - \frac{1}{L_f C_2} + \frac{R_f^2}{L_f^2}$; $T_7 = \frac{1}{k_{10}}(-k_9 k_{d1} T_1 + \frac{k_9}{R_1 C_1} + \frac{k_{d1}}{L_f} + \frac{k_9 Rd_1}{L_1} - 1)$; $T_8 = \frac{1}{k_{10}}(-k_9 k_{d1} T_2 - \frac{k_9}{C_1} - Rd_1)$; $T_9 = \frac{1}{k_{10}}(-k_9 k_{d1} T_3 - \frac{k_{d1}}{L_f})$; $T_{10} = \frac{1}{k_{10}}(-k_9 k_{d1} T_4)$; $T_{11} = \frac{1}{k_{10}}(-k_9 k_{d1} T_5)$; $T_{12} = \frac{1}{k_{10}}(-k_9 k_{d1} T_6 - \frac{k_9}{C_1} + \frac{k_{d1} R_f}{L_f})$; $T_{13} = \frac{D_2{}'}{L_2 R_2 C_2}$; $T_{14} = \frac{-D_2{}'^2}{L_2 C_2}$; $T_{15} = \frac{D_2{}' I_2}{L_2 C_2 V_M} - \frac{V_2}{L_2 V_M k_3}$; $T_{16} = \frac{V_2 k_2}{L_2 V_M k_3}$; $T_{17} = \frac{V_2 k_1}{L_2 V_M k_3}$; $T_{18} = \frac{\frac{k_4}{R_2 C_2} - k_{d2} k_4 T_{13} + \frac{(k_{d2} + k_4 Rd_2) D_2{}'}{L_2} - 1}{k_5 + k_{d2} k_4 T_{17}}$; $T_{19} = \frac{\frac{-k_4 D_2{}'}{C_2} - k_{d2} k_4 T_{14} - Rd_2}{k_5 + k_{d2} k_4 T_{17}}$; $T_{20} = \frac{\frac{k_4 I_2}{C_2 V_M} - k_{d2} k_4 T_{15} - \frac{(k_{d2} + k_4 Rd_2) V_2}{L_2 V_M}}{k_5 + k_{d2} k_4 T_{17}}$; $T_{21} = \frac{-k_{d2} k_4 T_{16}}{k_5 + k_{d2} k_4 T_{17}}$; $T_{160} = \frac{D_2{}'}{L_2 C_2}$; $T_{22} = \frac{-k_4 k_{d2} T_{160} + \frac{k_4}{C_2}}{T_{23}}$; $T_{23} = k_5 + k_4 k_{d2} T_{17}$; $T_{24} = \frac{-k_{d2} - k_4 Rd_2}{L_2 T_{23}}$; $T_{25} = \frac{\frac{k_4 k_{d2} D_2{}'}{L_2 C_2} - \frac{k_4}{C_2}}{T_{23}}$;

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:47:07 UTC from IEEE Xplore. Restrictions apply.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/2631-3871.htm

Convenience store locations prioritization: a fuzzy TOPSIS-GRA hybrid approach

Jyotdeep Singh
Mechanical Engineering, Delhi Technological University, Delhi, India
Parnika Tyagi
Electronics and Communication, Delhi Technological University, Delhi, India, and
Girish Kumar and Saurabh Agrawal
Mechanical Engineering, Delhi Technological University, Delhi, India

Received 29 January 2020
Revised 5 May 2020
31 August 2020
Accepted 22 October 2020

Abstract

Purpose – The objective of the study is to develop a methodology to strategically rank store locations using criteria such as population, store site characteristics, economic considerations, competition and so on to select the most optimal retail convenience store location.

Design/methodology/approach – A case of National Capital Region, India, for a 24-h convenience store was considered for the study and the major criteria that affect the performance of a convenience store are identified, such as population characteristics, economic criteria, competition, consumer accessibility, store size, total cost, site attractiveness and security. Fuzzy AHP is utilized to find the weightage for each criteria and a combination of fuzzy TOPSIS and grey relational analysis (GRA) is applied to rank the alternative using these criteria weight. Further, results obtained are compared with results from fuzzy TOPSIS and fuzzy VIKOR methods. Sensitivity analysis is also performed for ensuring the robustness of the framework.

Findings – It is observed that outcomes do not change under various settling coefficient values, demonstrating that the methodology is very robust. The developed framework will be quite useful to diverse retailers looking to expand and generate substantial profits.

Research limitations/implications – A large sample size of number of locations encourages generalization of results. Strategic ranking of the selected locations is carried out on a few selected criteria. The study was limited by the designated geographical area.

Originality/value – The study contributes to the few available articles on convenience store selection using combination of fuzzy AHP, fuzzy TOPSIS and GRA for a developing country.

Keywords Retailing strategy, Convenience store, AHP, TOPSIS, Grey relational analysis

Paper type Research paper

1. Introduction

The choice of retail location areas is an important decision for retailing decision-makers (DMs) as it is a long-term decision and higher costs are involved in it. As convenience is of utmost importance to the present buyers, a retail store can thrive or flop exclusively dependent on its location. The reason is that a decent area can pull in and be open to a substantial number of individuals affecting deals and benefits. Subsequently, it is hard to make up for the negative impact of an ill-placed location. A large chunk of convenience store profit depends on how "convenient" it is, whether that convenience comes from the layout, variety, brand or location.

© Jyotdeep Singh, Parnika Tyagi, Girish Kumar and Saurabh Agrawal. Published in *Modern Supply Chain Research and Applications*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at http://creativecommons.org/licences/by/4.0/legalcode

Modern Supply Chain Research and Applications
Emerald Publishing Limited
2631-3871
DOI 10.1108/MSCRA-01-2020-0001

A convenience store is a small retail business that stocks a variety of ordinary items, such as perishables, baby products, desserts, soda pops, common medicines, toiletries, magazines, papers and so on. These are different from general stores and town shops as they do not fall in a rural area and are better alternative to bigger stores. In all the retailing joint frameworks, convenience stores (CVs) have the easiest access to customers. Proper site determination entails much more than simply choosing a site with more visibility and ease of access. Successful research on rivalry, socioeconomics and advertise request are additionally required. Location is considered a prime consideration for retailing and services (Jones *et al.*, 2003). Hence, it becomes significant to give proper thought to decide on an appropriate location to increase foot traffic and revenue.

The motivation for the study comes from the increased interest in the convenience retailing space in India. The trend is likely to gather pace with more players competing to create a dominance in market. With change in time, there has been a slow and gradual change in shopping trends especially for household goods considering criteria such as good quality, good assortment and competitive prices. Owing to these changes there has been a 14.2% CAGR (compound annual growth rate) increase per year in organized retail sector (PWC, 2017). This motivated multiple brands to enter in convenience retailing business where previously over 90% market was controlled by unorganized retail shops. Further, India has the most astounding utilization development among the main ten nations positioned by size of family last utilization consumption (HFCE). Moreover, rising urbanization (27% towards the start of century to 32% in 2015) and the developing interest of ladies in the workforce are additionally making a period starved purchaser bunch which is probably going to pay a premium for comfort segment. One other purpose behind this change is the ongoing endorsement of Model Shops and Establishment (Regulation of Employment and Condition of Services) Bill. This will provide a legal framework to work 365 days a year and adaptable long periods of activity subsequently supporting the opening of nonstop accommodation stores. Despite the popularity and high growth estimates, there has been limited research done in this domain. Hence, for this research we consider the case of 24×7 open convenience stores as they are gaining popularity due to their "convenient" nature and due to our fast-paced life.

The research questions that have been addressed in this study are as follows:

(1) Selecting important criteria for convenience store location selection.

(2) Strategically ranking these selected criteria through a suitable technique.

(3) Developing a methodology to strategically rank store locations using criteria such as population, store, site characteristics, economic considerations, competition and so on.

(4) Selecting the most optimal retail convenience store location.

India's organized retail sector market is expected to grow by 83% by 2021 (Chakravorty, 2019), and convenience stores constitute a majority in it. Thus, selecting the optimal convenience store location becomes a priority in the business strategy of a firm when it is undergoing an expansionary phase whether locally or globally. A convenience store location is selected keeping in mind various criteria/core capabilities that determine its success; in this study, the authors attempt to optimize the selection process by making it more relevant to business owner's needs and the necessities of a developing country. Our study provides a low-cost, easily integrable, practical method to the strategy of industry professionals to grow their business and obtain a competitive advantage.

For the purpose of this research, the authors have studied the retail sector in India and the major opportunities that exist in this sector. The focus of this study is about the location

selection. Developing a convenience store location ranking methodology in the Indian retail context on the basis of various criteria specific to the sector and region under study is the main contribution of this study

A number of researchers have worked with area assessment for convenience stores. However, there is limited work in Indian context for the determination of convenience store locations. The fact that this study focusses on round-the-clock convenience stores also enhances the originality of this research. In the past, researchers have focussed more on exact estimations rather than the fuzzy environment. There are numerous criteria which should be assessed when choosing where to locate a convenience store. For instance, one of the assessment criteria depends on the population characteristics residing in the chosen territory. Besides, it being difficult to assess human behaviour and model characteristics, socioeconomic factors and behavioural aspects may also not be uniform over an area. Distinguishing these assessment criteria, characterizing the impacts of them on one another and their significance and picking an area require a very much characterized multi-criteria decision-making (MCDM)-based assessment (Burnaz and Topcu, 2006). Apart from the criteria considered in the literature, security is considered as a criterion affecting the location feasibility in the proposed study.

The rest of the paper is arranged as follows. The literature review is discussed in Section 2. The theoretic description of fuzzy analytical hierarchical process (AHP), fuzzy TOPSIS and grey relational analysis (GRA) is provided in Section 3. In section 4, a real-world case study is presented; the results and sensitivity analysis are discussed in Section 5. Finally, the paper is concluded in Section 6.

2. Literature review

This section deals with the detailed literature review under two categories; location selection methodologies and factors affecting retail location.

2.1 Location selection methods

In recent years, many traditional approaches have been used to solve various MCDM problems. Yang *et al.* (2008) used AHP (analytical hierarchical method)/ANP (analytical network process) approach to solve a MCDM problem of selecting best facility; however, the research adopted a non-fuzzy methodology and a heuristic approach for alternatives ranking. Chang and Hsieh (2014) used a TOPSIS model to determine best chain store location using five criteria: crowds, store cluster, site features, store site acreages and the proportion of rent expenses against annual sales. Gupta *et al.* (2018) applied MCDM tools for optimum locations identification for charging of electric vehicles. Bai and Satir (2020) suggested an integrated model using AHP, information entropy theory (IET) and entropy TOPSIS techniques to rate green supplier satisfaction. However, both the researches have adopted a non-fuzzy approach and focussed on limited criteria for deciding best location. Yap *et al.* (2019) reviewed number of articles for site selection using MCDM.

There were some studies for location selection under fuzzy environment. Kumar *et al.* (2019) employed fuzzy TOPSIS approach for preferential selection of locations for installing CCTV surveillance system. Singh *et al.* (2018) applied a fuzzy AHP approach to locate a warehouse in Iranian market; however, the study is limited by focussing only on four major criteria affecting a warehouse location.

The researchers started using integrated approach to exploit the advantages of more than two combined techniques. Gupta *et al.* (2019) proposed an integrated MCDM approach for green supplier selection using integrated fuzzy AHP and three other, MABAC, WASPAS and TOPSIS. Önüt *et al.* (2010) developed a combined fuzzy AHP for assigning weights and fuzzy

TOPSIS approach to locate a shopping store adopting a real-life case of Istanbul, Turkey, driven by major economic shifts in the country. Consequently, Nyaoga *et al.* (2016) proposed a combined fuzzy TOPSIS-GRA approach to evaluate value chain performance of tea processing chains, the study also focusses on ranking them based on certain pre-defined qualitative criteria and subsequently determining the ranks of performance of value chain forms.

However, several other unconventional methods have also been adopted for optimal location selection. Kaufmann *et al.* (2000) compared three multi-unit retail location site determination techniques, in particular a successive methodology, a worldwide methodology and a proposed foreseen defer strategy. The authors assessed the after effects of these site choice strategies by utilizing diversion like two-party re-enactment. Malik *et al.* (2015) used graph theory and matrix approach for selection of locations of collection centres for reverse logistics. Arsovski *et al.* (2017) proposed a model for location selection for parking lots based on fuzzy AHP and Hurwitz methods. Recently, Shan *et al.* (2019) developed an algorithm for the optimization of selection of facility location of chain stores in China. Kuo *et al.* (2002) developed a decision-making system for selecting the best location out of multiple alternatives using an integrated model of fuzzy AHP and artificial neural network. However, model training might be difficult to tackle where there is no data availability. Roig-Tierno *et al.* (2013) worked on a methodology for location selection for retail stores using geographic information system while integrating it with AHP. Mendes and Themido (2004) discussed use of genetic algorithms for the global problem of the multi-outlet chain configuration.

Lack of data and existence of several conflicting qualitative and quantitative factors determine best alternative location. The aim was to obtain a better relationship between criteria and optimally ranking the alternatives. Hence, owing to the evident superiority of integrated methods, a methodology which combined TOPSIS method with GRA was adopted by the authors. AHP is utilized to ascertain loads for various criteria and a hybrid fuzzy TOPSIS-GRA model is utilized to rank the elective areas. Although there are numerous examples of location selection based on the MCDM approach in the literature, there is no evidence of employing the hybrid fuzzy TOPSIS-GRA method in the field of location selection. The proposed model's results have been compared with the prominent and most used MCDM tools such as fuzzy TOPSIS and fuzzy VIKOR and also the model has been subjected to sensitivity analysis. This research contributes to the scarce literature of convenience store selection using combination of fuzzy AHP, fuzzy TOPSIS and GRA for a developing country.

2.2 Factors affecting convenience store locations

As per real-life examples and based on relevant literature and expert opinion, selection of a convenience store location has been deeply influenced by certain criteria and identifying these criteria is the first major step in convenience store location selection process. For selecting the most impactful criteria, a social, economic and geographical landscape has to be understood. Several of these factors have been mentioned by Kuo *et al.* (2002) that considered criteria such as competition, store characteristics and population characteristics for the selection of convenience store location. Furthermore, Burnaz and Topcu (2006) contemplated accessibility and competition factors in order to select the best location to open a clothing store. Önüt *et al.* (2010), in their study for locating a shopping store in Istanbul, Turkey, shortlisted population characteristics, degree of competition, accessibility, total cost, attractiveness as criteria affecting optimum location selection. Chang and Hsieh (2014) investigated the best chain store location using selection criteria's such as site features, the proportion of rent expenses against annual sales crowds, store cluster, to name a few. Erbıyık *et al.* (2012) used AHP for retail store site selection using shortlisted criteria such as costs,

competition conditions, traffic density and physical features. Turhan *et al.* (2013) provided a comprehensive overview on selecting retail store location based on certain factors that impact a store's performance and revenue. Ge *et al.* (2019) examined optimized superstore site selection incorporating measures such as population density, interest points, distribution of competitors, network of stores, transportations and so on. The authors employed a combination of machine learning and optimization methods for an optimized site selection. As per the expert opinion, security is a major concern for opening a convenience store in developing and underdeveloped countries. As a real-life example: It is well documented that 7-Eleven, a popular round-the-clock franchise, constantly adapts to the population characteristics, economic characteristics of the local market while setting up (LBS, 2013).

For this study, eight criteria are determined for the selection of the best location for opening a round-the-clock convenience store on the basis of the literature review and experts opinion. Refer Table 1 for the literature source for individual criteria. These criteria are: population characteristics, economic considerations, competition, consumer accessibility, store characteristics, total cost, site attractiveness and site security which are denoted as C_1, C_2, C_3, C_4, C_5, C_6, C_7, and C_8 respectively. A brief description of these criteria is provided here.

Population characteristics: Any business can't survive without customers and same is true for convenience stores. Thus, they are usually located in areas with population characteristics such as: population density, average age, gender demographics are favourable.

Economic criteria: The main factor that distinguishes convenience stores from unorganized retail stores is the cost and quality of items available for purchase along with additional services such as fast food, ATMs and so on. Thus, economic demographics such as average household income, purchasing power of the population in the trade area of convenience store play a vital role in selection of its location.

Competition: A Convenience store faces competition in the form of unorganized retail store, other convenience stores, big chain food marts and so on. The amount of competition in the trade area of a convenience store directly affects the amount of business it does.

Consumer accessibility: The easiness by which a consumer can visit a convenience store and shop is measured in the form of consumer accessibility, and it is affected by various factors such as road width, available parking spaces and so on. The easier a convenience store to shop, the more business it does.

Store size: The size of a convenience store determines the variety of goods and services it can provide. The size also affects the consumer shopping experience, thus directly affecting its business.

Total cost: The sum of cost associated with a location whether it's leased, bought or rented is called total cost. It is a vital criterion while selecting a location for a CVS as it has a direct effect on the total profit of the business.

Codes	Criteria	References
C_1	Population characteristics	Ingene and Brown (1987), Hoch *et al.* (1995), Kuo *et al.* (2002)
C_2	Economic criteria	Cottrell (1973), Irwin (1986), Yang *et al.* (2008), Berman and Evans (1995)
C_3	Competition	Craig (1984), Hoch *et al.* (1995), Kuo *et al.* (2002)
C_4	Consumer accessibility	Irwin (1986), Kuo *et al.* (2002), Yang *et al.* (2008), Erbıyık *et al.* (2012)
C_5	Store size	Irwin (1986), Karande and Lombard (2005), Erbıyık *et al.* (2012)
C_6	Total cost	Irwin (1986), Tzeng *et al.* (2002), Erbıyık *et al.* (2012)
C_7	Site attractiveness	Irwin (1986), Kuo *et al.* (2002), Erbıyık *et al.* (2012)
C_8	Security	Based on discussion with experts

Table 1. Factors affecting convenience store locations

Site attractiveness: It is not always necessary that the consumers solely consist of the population residing in the trade area. People from different regions might also act as consumers at times more so when there is a presence of a magnet such as a tourist spots, government buildings, corporate offices and so on.

Security: 24 × 7 functionality is one of the features of a convenience store which also differentiates it from the unorganized retail sector, but this time benefit is only possible if the there is sufficient security such as CCTV cameras, regular police patrols and so on. Security also affects the participation of the female gender as both employees and consumers.

3. Methodology

Retail site location selection is a tactical assessment. Therefore, for the selection of the most optimum location of the convenience store, MCDM tools have to be taken into consideration. Generally, in MCDM methods the decision-making process is done in the presence of large number of criteria or attributes.

In this research, a hybrid method of fuzzy TOPSIS-GRA is used because fuzzy TOPSIS as an independent approach requires subjective decision-making while weighing of various criteria (Olson, 2004). This major drawback puts hybrid approach of fuzzy AHP and fuzzy TOPSIS-GRA ahead of fuzzy TOPSIS. Furthermore, the traditionally weighted TOPSIS method only calculates the relative distance to the ideal solution, but ignores the curve trend and cannot accurately reflect the reality (Chen 2019). GRA is a measure of the non-linear relationship between sequences and can compensate for the deficiencies of TOPSIS method (Chen 2019). Hence, this paper uses a combined fuzzy TOPSIS-GRA approach, to evaluate and select the best location alternative for opening a convenience store. It then proceeds to verify the feasibility and effectiveness of the model by a series of sensitivity analysis.

The proposed methodology described in this section consists of two parts. First, fuzzy AHP is applied to calculate criteria weights. Next, a hybrid approach of fuzzy TOPSIS-GRA is employed to rank and select the best location alternative. A flowchart for the developed methodological framework for this study is shown in Figure 1. Steps involved in the methodology are detailed in the following subsections.

3.1 Fuzzy AHP

In suggested approach, fuzzy AHP is used to get the weights of the shortlisted criteria for choice of the best area to open a convenience store. AHP is an MCDM technique, which is utilized broadly for prioritizing issues in various fields, for example, legislative issues, financial aspects, social and the management issues (Işıklar and Büyüközkan, 2007). Pure AHP has a few deficiencies. The lack of quality related with the mapping of one's discernment or decisions to a number is not considered (Büyüközkan, 2004). AHP works just with accurate and conventional information, while researchers, as a result of their great capacity for subjective information preparing, settle on choices in fuzzy situations (Torfi *et al.*, 2009). The fuzzy augmentation of AHP, when utilized as chief, is generally discovered more advantageous to express interim decisions than fixed esteem decisions because of the fuzzy idea of the examination procedure. A substantial number of studies have utilized AHP incorporated with fuzzy rationale to take care of different issues. Singh *et al.* (2018) used fuzzy AHP for selection of warehouse location. Büyüközkan (2004) selected best software development strategy using this method.

3.2 Hybrid fuzzy TOPSIS and GRA approach

In this section, an integrated fuzzy MCDM approach is exhibited to assess the engaging quality of different convenience store location options. This strategy has been and can be

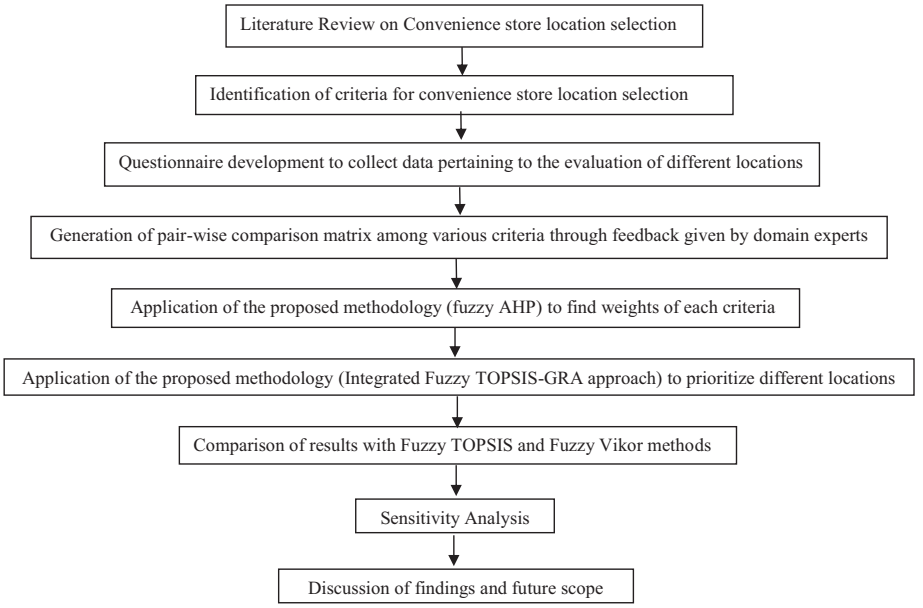


Figure 1.
Research framework

utilized to take care of issues containing uncertain, inconclusive and abstract information. A triangular fuzzy number is utilized to express the locations assessment on different criteria and furthermore locations assessment on choices regarding every paradigm. The membership function of a triangular fuzzy number $\tilde{A} = (j,\ k,\ l)$ is defined as per Eq. (1) (see Figure 2).

$$f_{\tilde{A}}(\mu) = \left\{0,\ \mu < j;\ \frac{\mu - j}{j - k},\ j \leq \mu \leq k;\ \frac{l - \mu}{l - k},\ k \leq \mu \leq l;\ 0,\ \mu > k\right\} \tag{1}$$

Algebraic operations (addition, subtraction, multiplication, division and reciprocal) on two triangular fuzzy numbers $\tilde{A} = (j_1,\ k_1,\ l_1)$ and $\tilde{B} = (j_2,\ k_2,\ l_2)$ (Kaufmann and Gupta 1991; Dubois and Prade 1980) are expressed as:

$$\tilde{A} + \tilde{B} \approx (j_1 + j_2,\ k_1 + k_2,\ l_1 + l_2) \tag{2}$$

$$\tilde{A} - \tilde{B} \approx (j_1 - j_2,\ k_1 - k_2,\ l_1 - l_2) \tag{3}$$

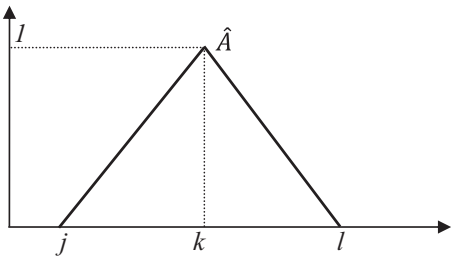


Figure 2.
Triangular fuzzy number

$$\tilde{A}.\tilde{B} \approx (j_1.\ j_2,\ k_1.\ k_2,\ l_1.\ l_2) \tag{4}$$

$$\tilde{A}/\tilde{B} \approx (j_1/j_2,\ k_1/k_2,\ \ l_1/l_2\ ) \tag{5}$$

$$1/\tilde{A} \approx (1/l_1,\ 1/k_1,\ 1/j_1\ ) \tag{6}$$

In the proposed strategy, the criteria loads are produced utilizing fuzzy AHP. The positioning of the location options is finished by the integrated fuzzy TOPSIS and GRA strategy. The steps (1–3) utilized Buckley's fuzzy AHP calculation (Buckley, 1985), and the steps (4–12) of the proposed integrated fuzzy TOPSIS and GRA calculation (Chang and Hsieh, 2014) are condensed as follows:

Step 1: A $p \times p$ pair-wise comparison matrix of criteria is created using linguistic variables. Respective fuzzy scale is then replaced with these variables as depicted in the matrix $\tilde{P}$.

$$\tilde{P} = \left(1\ \tilde{p}_{12\ldots\ldots}\ \tilde{p}_{1p} \vdots \ddots \vdots \tilde{p}_{p1}\ p_{p2}\ 1\ \right) \tag{7}$$

Where,

$$p_{ij} = \left\{ \left(\frac{1}{a_i},\ \frac{1}{b_i},\ \frac{1}{c_i}\right),\ \text{for } \forall\ i\ > j;\ (1,\ 1,\ 1) \text{ for } \forall\ i = j \text{ and } (a_i,\ \ b_i,\ c_i);\ \text{ for } \forall\ i < j \right\}$$

In the pair-wise criteria comparison matrix $\tilde{P}$, all the diagonal elements are 1 and among non-diagonal elements, the symmetric elements are reciprocal of each other.

Step 2: Geometric mean technique is used to calculate geometric mean of individual row of $\tilde{P}$ and subsequently a matrix of fuzzy relative weight is constructed using equation (8) (Chang and Hsieh, 2014).

$$\tilde{r}_j = \left[\prod_{l=1}^{p} \tilde{r}_{jl}\right]^{1/p} \text{for} j,\ l = 1, 2,\ \ldots,\ p \tag{8}$$

Step 3: Consequently, the fuzzy weights of each criterion are calculated using Eq. (9)

$$\tilde{w}_j\ \ = \tilde{r}_j \otimes \left[\prod_{j=1}^{p} \tilde{r}_{jl}\right]^{-1} \tag{9}$$

Where the fuzzy variable $\tilde{w}_j$ is the fuzzy weight for the jth criterion which is the ratio of jth criteria geometric mean $\tilde{r}_j$ and summation of geometric mean column. Being a fuzzy number, it is specified as $\tilde{w}_j = (w_{j1},\ w_{j2},\ \ w_{j3})$. Fuzzy weight for the jth criterion has lower, middle and upper values represented as $(w_{j1},\ w_{j2},\ w_{j3})$.

Step 4: A panel of DMs who are expert in retail and convenience store site selection process is selected. A decision matrix consisting of the evaluations of each alternative with respect to each of the selected criteria by a panel of DMs is constructed. For this study, the linguistic variables chosen for this purpose and their corresponding triangular fuzzy numbers are shown in Table 2. The decision matrix for "q" number of alternate locations and "p" number of criteria is expressed as Eq. (10).

$$\tilde{A} = [\tilde{x}_{11}\ \tilde{x}_{12}\ \tilde{x}_{13}\ \ldots \tilde{x}_{q1}\ \tilde{x}_{q2}\ \tilde{x}_{qp}\ ] = [\tilde{x}_{ij}]_{q \times p} \tag{10}$$

Where $\tilde{x}_{ij} = \frac{1}{K}[\tilde{x}_{ij}^1 + \tilde{x}_{ij}^2 + \cdots + \tilde{x}_{ij}^K] = \frac{1}{K}\sum_{e=1}^{K} \tilde{x}_{ij}^e$ that is, it is average of k expert rating as all experts are given equal weights.

Where performance rating of alternative A_i for criterion C_j is $\tilde{x}_{ij}; i = 1, 2, \ldots, q$; $j = 1, 2, \ldots, p$. When it is assessed by kth expert, $\tilde{x}_{ij}$ is defined as $(m_{ij}^k, n_{ij}^k, p_{ij}^k)$, a triangular fuzzy number.

Step 5: Normalization of the fuzzy decision matrix is established on the traits of two types of criteria which are smaller-the-better (cost) criteria and larger-the-better (benefit) respectively.

For larger the better criteria, normalization is done using Eq. (11).

$$\frac{\tilde{r}_{ij} = \tilde{x}_{ij} - \min(\tilde{x}_{ij})}{(\tilde{x}_{ij}) - \min(\tilde{x}_{ij})} \tag{11}$$

Where $\tilde{x}_{ij}$ is the element of decision matrix A_i obtained in previous step.

For smaller the better criteria, normalization is done using Eq. (12).

$$\frac{\tilde{r}_{ij} = \max(\tilde{x}_{ij}) - \tilde{x}_{ij}}{(\tilde{x}_{ij}) - \min(\tilde{x}_{ij})} \tag{12}$$

Step 6: The multiplication operation between the normalized decision matrix and the criteria weights gives the weighted decision matrix in the normalized form. The weighted normalized decision matrix $\tilde{V}$ for individual criterion is expressed by Eq. (13).

$$\tilde{V} = [\tilde{v}_{ij}]_{p \times q} \text{ for } \quad i = 1, 2, 3, \ldots p; \quad j = 1, 2, \ldots, q \tag{13}$$

where, $\tilde{v}_{ij} = \tilde{r}_{ij} \times \tilde{w}_j$ and being a fuzzy number, it is expressed as

		C1	C2	C3	C4	C5	C6	C7	C8
L_1	DM1	(7, 9, 10)	(5, 7, 9)	(7, 9, 10)	(5, 7, 9)	(5, 7, 9)	(7, 9, 10)	(5, 7, 9)	(5, 7, 9)
	DM2	(5, 7, 9)	(5, 7, 9)	(5, 7, 9)	(7, 9, 10)	(3, 5, 7)	(7, 9, 10)	(5, 7, 9)	(7, 9, 10)
	DM3	(5, 7, 9)	(7, 9, 10)	(7, 9, 10)	(7, 9, 10)	(7, 9, 10)	(5, 7, 9)	(7, 9, 10)	(5, 7, 9)
	DM4	(7, 9, 10)	(5, 7, 9)	(7, 9, 10)	(7, 9, 10)	(5, 7, 9)	(5, 7, 9)	(3, 5, 7)	(5, 7, 9)
L_2	DM1	(5, 7, 9)	(3, 5, 7)	(0, 1, 3)	(3, 5, 7)	(5, 7, 9)	(3, 5, 7)	(1, 3, 5)	(0, 1, 3)
	DM2	(3, 5, 7)	(3, 5, 7)	(1, 3, 5)	(1, 3, 5)	(3, 5, 7)	(1, 3, 5)	(1, 3, 5)	(1, 3, 5)
	DM3	(3, 5, 7)	(5, 7, 9)	(0, 1, 3)	(1, 3, 5)	(3, 5, 7)	(3, 5, 7)	(0, 1, 3)	(1, 3, 5)
	DM4	(5, 7, 9)	(7, 9, 10)	(0, 1, 3)	(0, 1, 3)	(3, 5, 7)	(3, 5, 7)	(3, 5, 7)	(0, 1, 3)
L_3	DM1	(3, 5, 7)	(3, 5, 7)	(1, 3, 5)	(3, 5, 7)	(0, 1, 3)	(5, 7, 9)	(5, 7, 9)	(3, 5, 7)
	DM2	(3, 5, 7)	(1, 3, 5)	(3, 5, 7)	(5, 7, 9)	(3, 5, 7)	(7, 9, 10)	(5, 7, 9)	(1, 3, 5)
	DM3	(5, 7, 9)	(3, 5, 7)	(1, 3, 5)	(1, 3, 5)	(1, 3, 5)	(3, 5, 7)	(7, 9, 10)	(3, 5, 7)
	DM4	(1, 3, 5)	(5, 7, 9)	(3, 5, 7)	(5, 7, 9)	(1, 3, 5)	(5, 7, 9)	(5, 7, 9)	(5, 7, 9)
L_4	DM1	(1, 3, 5)	(3, 5, 7)	(0, 1, 3)	(1, 3, 5)	(3, 5, 7)	(1, 3, 5)	(5, 7, 9)	(1, 3, 5)
	DM2	(1, 3, 5)	(0, 1, 3)	(0, 1, 3)	(1, 3, 5)	(1, 3, 5)	(1, 3, 5)	(5, 7, 9)	(3, 5, 7)
	DM3	(3, 5, 7)	(1, 3, 5)	(0, 1, 3)	(5, 7, 9)	(3, 5, 7)	(0, 1, 3)	(3, 5, 7)	(5, 7, 9)
	DM4	(0, 1, 3)	(3, 5, 7)	(1, 3, 5)	(3, 5, 7)	(3, 5, 7)	(1, 3, 5)	(5, 7, 9)	(5, 7, 9)
L_5	DM1	(5, 7, 9)	(1, 3, 5)	(3, 5, 7)	(5, 7, 9)	(5, 7, 9)	(1, 3, 5)	(5, 7, 9)	(3, 5, 7)
	DM2	(3, 5, 7)	(0, 1, 3)	(1, 3, 5)	(3, 5, 7)	(3, 5, 7)	(3, 5, 7)	(7, 9, 10)	(1, 3, 5)
	DM3	(5, 7, 9)	(1, 3, 5)	(1, 3, 5)	(3, 5, 7)	(3, 5, 7)	(5, 7, 9)	(3, 5, 7)	(1, 3, 5)
	DM4	(7, 9, 10)	(3, 5, 7)	(5, 7, 9)	(1, 3, 5)	(3, 5, 7)	(1, 3, 5)	(7, 9, 10)	(0, 1, 3)
L_6	DM1	(3, 5, 7)	(1, 3, 5)	(0, 1, 3)	(1, 3, 5)	(3, 5, 7)	(3, 5, 7)	(7, 9, 10)	(1, 3, 5)
	DM2	(5, 7, 9)	(3, 5, 7)	(0, 1, 3)	(5, 7, 9)	(3, 5, 7)	(5, 7, 9)	(5, 7, 9)	(0, 1, 3)
	DM3	(3, 5, 7)	(0, 1, 3)	(1, 3, 5)	(3, 5, 7)	(1, 3, 5)	(1, 3, 5)	(7, 9, 10)	(0, 1, 3)
	DM4	(5, 7, 9)	(3, 5, 7)	(3, 5, 7)	(3, 5, 7)	(5, 7, 9)	(3, 5, 7)	(7, 9, 10)	(0, 1, 3)

Table 2. Responses of decision-makers

$\tilde{v}_{ij} = (m^*_{ij},\ n^*_{ij},\ p^*_{ij}),\ i = 1, 2, 3, \ldots q;\quad j = 1, 2, \ldots, p;$ For q alternates and p criteria

Step 7: Fuzzy positive ideal solution (FPIS) and fuzzy negative solution (FNIS) are obtained from the weighted normalized decision matrix $\tilde{V}$ using Eqn (14) and (15) respectively.

$$\tilde{V}^+ = \left(\tilde{v}_j^+\right) = \left(\tilde{v}_1^+,\ \tilde{v}_2^+,\ \ldots,\ \tilde{v}_p^+\right), j = 1, 2, 3, \ldots p \tag{14}$$

$$\tilde{V}^- = \left(\tilde{v}_1^-\right) = \left(\tilde{v}_1^-,\ \tilde{v}_2^-,\ \ldots,\ \tilde{v}_p^-\right), j = 1, 2, 3, \ldots p \tag{15}$$

where $v_j^- = (m_j^-,\ n_j^-,\ p_j^-)$ and $v_j^+ = (m_j^+,\ n_j^+,\ p_j^+),\ j = 1, 2, 3, \ldots p$ and these are the lowest values for FNIS and highest values for FPIS for a criterion in the given column.

Step 8: In this step, separation distances from FPIS and FNIS are calculated. There are various methods to calculate the separation measures $\tilde{V}_{j^*}^{d^+}$ from FPIS and $\tilde{V}_{j^*}^{d^-}$ FNIS of each alternative to resolve grey relational coefficient of every alternative. Some of these method are: normalized Hamming distance (l), normalized Euclidean distance (h), Hamming distance (d), Euclidean distance (c). This study is based on expansion of Burillo and Bustince's technique (1996) to calculate grey relational coefficient based on the Hamming distance (d). The separation measures $\tilde{V}_{j^*}^{d^+}$ from FPIS and $\tilde{V}_{j^*}^{d^-}$ NIS of each alternative are evaluated using Eqn. (11) and (12).

$$\tilde{V}_{j^*}^{d^+} = \frac{1}{4}\sum_{i=1}^{q}\left[\left|m^*_{ij} - m_i^+\right| + \left|n^*_{ij} - n_i^+\right| + \left|p^*_{ij} - p_i^+\right|\right] \tag{16}$$

$$\tilde{V}_{j^*}^{d^-} = \frac{1}{4}\sum_{i=1}^{q}\left[\left|m^*_{ij} - m_i^-\right| + \left|n^*_{ij} - n_i^-\right| + \left|p^*_{ij} - p_i^-\right|\right] \tag{17}$$

Step 9: The following equation is used to generate grey relational coefficient of individual alternative from FPIS and FNIS respectively.

$$\frac{\gamma^+\left(\tilde{V}_{j^*}^{d^+}\right) = \min_i\min_j\ \tilde{V}_{j^*}^{d^+} + \zeta\max_i\max_j\ \tilde{V}_{j^*}^{d^+}}{\tilde{V}_{j^*}^{d^+} + \zeta\max_i\max_j\ \tilde{V}_{j^*}^{d^+}} \tag{18}$$

Here, the ζ is resolving coefficient which is usually taken as 0.5

$$\frac{\gamma^-\left(\tilde{V}_{j^*}^{d^-}\right) = \min_i\min_j\ \tilde{V}_{j^*}^{d^-} + \zeta\max_i\max_j\ \tilde{V}_{j^*}^{d^-}}{\tilde{V}_{j^*}^{d^-} + \zeta\max_i\max_j\ \tilde{V}_{j^*}^{d^-}} \tag{19}$$

Step 10: Calculation of grey relational grade from grey relational coefficient of individual alternative using Eqn. (20) and (21).

$$d_i^+ = \sum_{j=1}^{q}\gamma^+\left(\tilde{V}_{j^*}^{d^+}\right) \tag{20}$$

$$d_i^- = \sum_{j=1}^{q} \gamma^- \left(\tilde{V}_{j^*}^{d^+} \right) \quad (21)$$

Step 11: The closeness coefficient CC_i defined as closeness to the ideal solution is calculated by the following expression.

$$CC_i = \frac{d_i^+}{d_i^-} \quad (22)$$

Step 12: The preference order is ranked. Using the index calculated in the previous step, a decreasing order is followed by the alternatives.

4. Case illustration

To illustrate fuzzy AHP–fuzzy TOPSIS–GRA hybrid approach suggested in this work, a case study for a nationwide round-the-clock convenience store chain has been taken.

4.1 Profile of the case company

This study has considered an Indian round-the-clock convenience store chain known as ABC which is anticipating extending its business in the National Capital Region (NCR) of Delhi through widening its market reach and nearness in Delhi and other close-by territories. ABC is India's sole composed retail chains in the "round the clock" accommodation store group propelled with the vision of turning into a vital piece of the lives of individuals in each network.

The premier ABC convenience store was opened in 2004. The company comprises of stores open 24 h every day, seven days a week. ABC store offers more than 3,500 items and administrations. Its contributions and format of stores are intended to take into account clients who are continually in a hurry. The New Delhi–based organization as of now works as ABC stores in the NCR and other northern areas of the world's second-most crowded nation. The organization has extension plan of ABC outlets in near future. The stores' main value proposition is to cater to customers who are constantly on the go, by being situated at close proximity to their homes and offices.

NCR of Delhi was chosen for this study, as the company plans to expand in New Delhi and the nearby areas such as Gurugram and so on. Delhi-NCR has been drawing in a huge workforce from smaller urban areas and rustic territories; it has caused an extensive increment in the number of inhabitants in the enormous urban areas of India. The development in populace in these metropolitan urban areas additionally has developed new spending interest areas. Rapid urbanization and higher disposable income have created new spending demand areas and customers have now started to choose convenience over price.

4.2 Data collection

Based on the review of relevant literature and the experience of a team of experts, eight criteria are determined for the selection of the best location for opening a convenience store. The team of experts comprises of four specialists in both the fields of academia and industry. The first two experts are professors in the field of operations research and supply chain management at premier management and engineering institutes in the country and have high-impact research in the field of MCDM. The rest two experts belong from the industry and are senior managers in the parent company chosen for this study. Detailed questionnaire was sent to the aforementioned experts. A lot of face-to-face interviews were

also held to determine the shortlisted criteria. The potential locations are Safdarjung Development Area (SDA) Market; Chattarpur Enclave; Prashant Vihar; Dwarka, Sector-22; Global Business Park, Gurgaon and Windsor IT Park, Noida which are denoted as L_1, L_2, L_3, L_4, L_5 and L_6, respectively. All these identified location sites belong to the National Capital Territory of Delhi region.

SDA Market, Safdarjung

(1) It is situated opposite to IIT Delhi and near NIFT, New Delhi.

(2) Situated in a busy market having wide variety of restaurants and clubs.

(3) Good public transport connectivity.

(4) High competition in the trade area.

(5) Good brand loyalty in the trade area.

Chattarpur Enclave

(1) Situated near one of the fanciest residential areas in South Delhi.

(2) Residents have high household monthly income and purchasing power.

(3) Very low degree of competition in the trade area.

(4) Good parking convenience.

(5) Great store visibility.

(6) Situated on main road.

(7) Security concerns near the site.

Prashant Vihar, Rohini

(1) Residential area nearby with high population density.

(2) High site attractiveness; surrounded by magnets such as cinemas, markets.

(3) Great public transport connectivity, metro station nearby.

(4) Low degree of competition in the trade area.

(5) Low brand loyalty.

Sector-22, Dwarka

(1) Residential area nearby with low land cost.

(2) High site attractiveness; surrounded by magnets such as schools and cinemas.

(3) Low degree of competition in the trade area.

(4) High population growth rate.

(5) Good public transport connectivity.

Global Business Park, Gurgaon

(1) High site attractiveness; surrounded by magnets such as corporate offices.

(2) High expected demand at night due to large number of MNCs in the trade area.

(3) Great public transport connectivity, metro station nearby.

(4) Low degree of competition in the trade area.

Sector-125, Noida

(1) It is situated near developing business centre with large number of MNCs.

(2) Very low degree of competition in the trade area.

(3) High site attractiveness; surrounded by magnets such as public parks, bird sanctuary and so on.

(4) High security concerns near the site.

(5) Good parking convenience.

4.3 Evaluation of criteria weights

The group of experts decides the overall significance of the criteria utilizing the idea of semantic variable and fuzzy AHP. The area choices are evaluated utilizing etymological factors. The pair-wise comparison matrix is constructed based on opinion of the four experts for following Eq. (7). Further, fuzzy relative weight matrix for all criteria is obtained based on Eqs. (8) and (9) and the results are shown in Table 3.

From Table 3, it is observed that the maximum priority weight achieved among all the different criteria is for the total cost followed by the population characteristics. This implies that to select the optimal convenience store location, the organizations should give a lot of emphasis on the total cost which comprises the ownership, renting plus maintained cost of that particular area/locality and also the population characteristics which include the population demographics such as age, gender, education levels and so on.

4.4 Ranking of alternatives

Four experts were chosen to assess the location options dependent on their expertise. All specialists were having over ten years of involvement in the area of retail and were functioning as senior researchers, senior managers and territory heads in the organization. At that point every individual expert assessed every one of the area selected based on the eight shortlisted criteria.

A fuzzy decision matrix is constructed using Eqs. (10), and it is shown in Table 2 for all the DMs $j = 1, 2, 3, 4$. The average fuzzy performance values are computed giving equal importance to all the DMs.

Further, the normalized fuzzy decision matrix is determined by utilizing Equations (11) and (12). The weighted fuzzy standardized matrix is determined by utilizing Eq. (12) concerning the fuzzy weights of the criteria which are acquired utilizing fuzzy AHP (Equations 7–9). Additionally, the fuzzy positive (FPIS) and fuzzy negative ideal solutions (FNIS) are evaluated by utilizing Equations (14) and (15), and these are given in Table 4.

Weight	Criteria
0.049, 0.141, 0.437	Population characteristics
0.025, 0.082, 0.213	Economic criteria
0.082, 0.111, 0.213	Competition
0.035, 0.107, 0.343	Consumer accessibility
0.016, 0.039, 0.142	Store size
0.052, 0.133, 0.451	Total cost
0.019, 0.072, 0.152	Site attractiveness
0.037, 0.116, 0.343	Security

Table 3.
Fuzzy weight matrix for criteria

Table 4.
FPIS and FNIS values of matrix

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
Average fuzzy decision matrix								
L_1	(6, 8, 9.5)	(5.5, 7.5, 9.25)	(6.5, 8.5, 9.75)	(6.5, 8.5, 9.75)	(5, 7, 8.75)	(6, 8, 9.5)	(5, 7, 8.75)	(5.5, 7.5, 9.25)
L_2	(4, 6, 8)	(4.5, 6.5, 8.25)	(0.25, 1.5, 3.5)	(1.25, 3, 5)	(3.5, 5.5, 7.5)	(2.5, 4.5, 6.5)	(1.25, 3, 5)	(0.5, 2, 4)
L_3	(3, 5, 7)	(3, 5, 7)	(2, 4, 6)	(3.5, 5.5, 7.5)	(1.25, 3, 5)	(5, 7, 8.75)	(5.5, 7.5, 9.25)	(3, 5, 7)
L_4	(1.25, 3, 5)	(1.73, 3.5, 5.5)	(0.25, 1.5, 3.5)	(2.5, 4.5, 6.5)	(2.5, 4.5, 6.5)	(0.75, 2.5, 4.5)	(4.5, 6.5, 8.5)	(3.5, 5.5, 7.5)
L_5	(5, 7, 8.75)	(1.25, 3, 5)	(2.5, 4.5, 6.5)	(3, 5, 7)	(3.5, 5.5, 7.5)	(2.5, 4.5, 6.5)	(5.5, 7.5, 9)	(1.25, 3, 5)
L_6	(4, 6, 8)	(1.75, 3.5, 5.5)	(1, 2.5, 4.5)	(3, 5, 7)	(3, 5, 7)	(3, 5, 7)	(6.5, 8.5, 9.75)	(0.25, 1.5, 3.5)
The weighted fuzzy normalized decision matrix								
L_1	(0.049, 0.141, 0.437)	(0.025, 0.082, 0.213)	(0.000, 0.000, 0.000)	(0.035, 0.107, 0.343)	(0.016, 0.039, 0.142)	(0.000, 0.000, 0.000)	(0.014, 0.052, 0.120)	(0.037, 0.116, 0.343)
L_2	(0.028, 0.085, 0.291)	(0.019, 0.064, 0.163)	(0.111, 0.310, 0.816)	(0.000, 0.000, 0.000)	(0.010, 0.024, 0.095)	(0.035, 0.085, 0.271)	(0.000, 0.000, 0.000)	(0.002, 0.010, 0.030)
L_3	(0.018, 0.057, 0.194)	(0.010, 0.037, 0.100)	(0.080, 0.199, 0.489)	(0.015, 0.049, 0.180)	(0.000, 0.000, 0.000)	(0.010, 0.024, 0.068)	(0.016, 0.059, 0.136)	(0.020, 0.067, 0.209)
L_4	(0.000, 0.000, 0.000)	(0.003, 0.009, 0.025)	(0.111, 0.310, 0.816)	(0.008, 0.029, 0.108)	(0.005, 0.015, 0.057)	(0.052, 0.133, 0.451)	(0.012, 0.046, 0.112)	(0.023, 0.077, 0.238)
L_5	(0.039, 0.039, 0.364)	(0.000, 0.000, 0.000)	(0.071, 0.177, 0.424)	(0.012, 0.039, 0.144)	(0.010, 0.024, 0.095)	(0.035, 0.085, 0.271)	(0.016, 0.059, 0.128)	(0.007, 0.029, 0.089)
L_6	(0.028, 0.029, 0.291)	(0.003, 0.009, 0.0250	(0.098, 0.266, 0.685)	(0.012, 0.039, 0.144)	(0.007, 0.019, 0.076)	(0.030, 0.072, 0.225)	(0.019, 0.072, 0.152)	(0.000, 0.000, 0.000)
Fuzzy PIS and NIS								
A^*	(0.049, 0.141, 0.437)	(0.025, 0.082, 0.213)	(0.111, 0.310, 0.816)	(0.035, 0.107, 0.343)	(0.016, 0.039, 0.142)	(0.052, 0.133, 0.451)	(0.019, 0.072, 0.152)	(0.037, 0.116, 0.343)
A^-	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)	(0.000, 0.000, 0.000)

Next, the separation measures $\tilde{V}_{j^*}^{d^+}$ from FPIS and $\tilde{V}_{j^*}^{d^-}$ FNIS of each alternative are evaluated using Eqn. (16) and (17). The grey relational coefficient of each alternative from FPIS and FNIS is calculated using Equations (18) and (19). Then, the grey relational grade is calculated (Equations 20–21), and these are shown in Table 5. Finally, the closeness coefficient using Eqn. (22) of each alternative is calculated and is shown in Table 6. Here, the resolving coefficient is taken to be $\zeta = 0.5$.

5. Results and discussion

The CC_i values of each location alternative L_1, L_2, L_3, L_4, L_5 and L_6 are found out to be 1.607, 0.919, 0.928, 1.097, 1.012 and 0.986 respectively. Therefore, the ranking order of the six location alternatives is $L_1 > L_4 > L_6 > L_5 > L_3 > L_2$. Thus, it is clear that SDA Market is the most appropriate location alternative to open a convenience store. It is followed by Sector-22 Dwarka, then by Global Business Park, Gurgaon; Windsor IT Park, Sector-125, Noida; Prashant Vihar and finally by Chattarpur Enclave.

The following reasons can be attributed for selecting SDA Market as the most appropriate location alternative to open a convenience store: site attractiveness (vicinity of popular universities such as IIT Delhi and NIFT, New Delhi, and it being a busy market having wide variety of restaurants and cafes which are open late into the night, thus increasing the demand for round-the-clock establishments), population characteristics (young crowd in need of essentials and food items round-the-clock), consumer accessibility (It has great public transport connectivity with metro stations being in close vicinity of the market), competition (There are no stores which solve the problem of round-the-clock convenience in the vicinity).

Based on our observations across different criteria and different locations, we can conclude that given the fact that the high cost of the SDA Market as a location alternative negatively affects the CC_i ranking of a location, while other favourable criteria override cost to generate positive results. It is observed from the discussion earlier that even though a location alternative is not the most optimal one in terms of individual attributes, the ranking comes out on the basis of the generation of a score based on the consideration of all the criteria considered in this study. Out of the six location alternatives, one of the location alternatives emerged as the clear winner but the rest with their respective scores are quite tightly scored demonstrating the highly sensitive nature of the location alternatives. This further implies that management authorities of the respective organizations should follow a logical and scientific approach for choosing the location of the convenience store. As a result of the study, we find that the proposed method is practical for ranking alternatives with respect to multiple conflicting criteria.

Furthermore, to authenticate the effectiveness and applicability of the methodology proposed in this study, a relative study is applied with other approaches. The proposed hybrid fuzzy TOPSIS–GRA methodology and the results obtained from it are compared with the results derived from two of the most widely used MCDM tools: fuzzy TOPSIS and fuzzy VIKOR.

In this research, a hybrid method of fuzzy TOPSIS–GRA is used because fuzzy TOPSIS as an independent approach requires subjective decision-making while weighing of various criteria (Olson, 2004). This major drawback puts hybrid approach of fuzzy AHP and fuzzy TOPSIS-GRA ahead of fuzzy TOPSIS. Furthermore, the traditionally weighted TOPSIS method only calculates the relative distance to the ideal solution, but ignores the curve trend and cannot accurately reflect the reality (Chen 2019). GRA is one of the best methods to make decisions under business environment. It is a measure of the non-linear relationship between sequences and can compensate for the deficiencies of TOPSIS method (Chen 2019). Hence, this paper uses a combined fuzzy TOPSIS–GRA approach, to evaluate and select the best location alternative for opening a convenience store.

Table 5.
Grey relational grade for alternatives

	PIS								NIS									
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	d^+	d^-
L_1	1	1	0.33	1	1	333	0.86	1	0.33	0.33	1	0.33	0.33	1	0.4	0.33	6.53	4.06
L_2	0.58	0.68	1	0.33	0.59	0.56	0.42	0.33	0.44	0.39	0.33	1	0.43	0.45	1	0.86	4.51	4.9
L_3	0.47	0.48	0.57	0.5	0.33	0.37	1	0.53	0.54	0.52	0.45	0.5	1	0.76	0.37	0.46	4.26	4.58
L_4	0.33	0.36	1	0.42	0.45	1	0.79	0.59	1	0.81	0.33	0.62	0.56	0.33	0.42	0.42	4.94	4.5
L_5	0.63	0.33	0.52	0.46	0.59	0.56	0.95	0.38	0.42	1	0.48	0.55	0.43	0.45	0.37	0.66	4.43	4.37
L_6	0.53	0.36	0.77	0.46	0.51	0.51	1	0.33	0.47	0.81	0.37	0.55	0.49	0.49	0.33	1	4.47	4.53

The ranking of TOPSIS-GRA method when compared with fuzzy TOPSIS and fuzzy VIKOR methods is very similar. The ranking in the proposed methodology and fuzzy TOPSIS is exactly the same, whereas when compared with fuzzy VIKOR, the rankings are similar except for the rankings of L_2, L_4, L_5 and L_6. The changes in rankings are due to the highly competitive nature of the locations selected. As there is no MCDM method which can guarantee the optimal solution in absolute terms, the comparison with one or more methods justifies the results of the proposed method. Further, sensitivity analysis validates the robustness of the suggested methodology.

Further, the resolving coefficient value is used for sensitivity analysis in order to verify the robustness of the proposed method. In the analysis, various resolving coefficient values ranging from 0.1 to 1 are used. It was observed that they do not affect the ranking order of the location alternatives. But as the ζ value increases, the CC_i values of alternatives converge. For the various ζ values Table 7 establishes the varied CC_i values that rely on the extension of Burillo and Bustince's method that is subject to the Hamming separation (d). As observed from the table, L_1 is the most ideal area selected as for closeness coefficient CC_i. The positioning request of the area choices does not change with the adjustments in the qualities. Lastly, the ranking of priority for allocation alternatives using the results of CC_i values is achieved as $L_1 > L_4 > L_6 > L_5 > L_3 > L_2$ with regard to assorted values of coefficient variable. Results of sensitivity analysis are shown in Table 7. Also, results are also presented graphically in Figure 3.

6. Conclusion

The urbanization of populace and the huge move of individuals towards comfort have made new shopping patterns which have brought about the rise of round-the-clock convenience stores. In the recent years, the convenience store portion advertisement has developed and has turned out to be exceedingly aggressive, with a fight to pick up the piece of the overall industry and offer of wallet of every single buyer. Choosing the most profitable site location turns into one of the most crucial and difficult undertakings for the organization.

In this paper, a fuzzy MCDM method is presented combining fuzzy TOPSIS and GRA methods, based on the extension of Burillo and Bustince's formula based on Hamming distance, in order to evaluate and select the best location alternative for opening a convenience store. A real-world case study from NCT of New Delhi was presented to explain the suggested approach. First, the criteria that make the location alternative more suitable to open a convenience store are determined, that is, population characteristics, economic considerations, competition, consumer accessibility, store characteristics, total cost, site attractiveness and site security.

Some examples of studies in similar domain are observed and compared with this study. The ranking of the criteria in this study is compared with a couple of studies in the retail

Table 6. Comparative results of different methodologies

Location alternatives	Proposed hybrid fuzzy TOPSIS–GRA methodology		Fuzzy TOPSIS		Fuzzy VIKOR	
	CC_i values	Rank	CC_i values	Rank	Q_j values	Rank
L_1	1.607	1	0.128588	1	0.027401901	1
L_2	0.919	6	0.092745	6	0.052781847	2
L_3	0.928	5	0.092943	5	0.066234786	5
L_4	1.097	2	0.12684	2	0.069690141	6
L_5	1.012	4	0.093856	4	0.053533376	3
L_6	0.986	3	0.098611	3	0.055429162	4

Table 7.
Sensitivity analysis results

Location	CC_i values									
	$\zeta = 0.1$	$\zeta = 0.2$	$\zeta = 0.3$	$\zeta = 0.4$	$\zeta = 0.5$	$\zeta = 0.6$	$\zeta = 0.7$	$\zeta = 0.8$	$\zeta = 0.9$	$\zeta = 1.0$
L_1	2.132	1.905	1.752	1.644	1.607	1.499	1.448	1.407	1.357	1.344
L_2	0.705	0.801	0.853	0.884	0.919	0.920	0.930	0.939	0.913	0.950
L_3	0.652	0.779	0.834	0.866	0.928	0.901	0.912	0.921	0.921	0.934
L_4	1.253	1.143	1.100	1.078	1.097	1.055	1.048	1.043	1.009	1.035
L_5	0.741	0.880	0.932	0.957	1.012	0.979	0.984	0.988	0.953	0.992
L_6	0.968	0.975	0.980	0.984	0.986	0.988	0.989	0.990	0.958	0.992

domain for developing economies. For example, in a similar study conducted in Taiwan aimed at selecting the most optimal convenience store location (Kuo *et al.*, 2002), "Competition" emerges as the top criterion, whereas in our study based in India, "Total Cost" incurred emerges as the deciding factor. The following reason can be attributed to this observation. As the market in Taiwan already has a highly competitive landscape and the number of players are quite higher when compared to the geography considered in this study. However in India, as round-the-clock convenience store is still a relatively newer concept with fewer players and less market share, hence "Total Cost" becomes a major deciding factor.

The second example considered is of a study conducted in Istanbul, Turkey, which aims to model shopping centre site selection (Önüt *et al.*, 2010). In their study, the deciding factor comes out to be "Total Cost", which is similar to our study. The second most important factor is "Economical", whereas in our study it is "Population characteristics". The reason for this observation can be attributed to the scale of investment in both the cases, as to set up a shopping centre investment is fairly high, hence the return on investment and estimated profits become a major attribute but to set up a convenience store the investment is much lower.

Additionally the main theoretical contribution and scientific value addition of this paper can be stated as follows:

(1) The first and foremost contribution of this study is usage of objective weight value as calculated by the fuzzy AHP to replace the subjective weight that DMs set directly that is, the objective weight of the fuzzy AHP is used instead of the subjective weight, which reduces the bias that may be caused by the subjective judgment.

(2) Second contribution is to integrate the TOPSIS method with GRA techniques in a fuzzy environment to assess and rank the location alternatives. The closeness coefficient is used as the evaluation basis for appropriate selection of the location alternative.

(3) Further, to check the heftiness and robustness of the created structure, sensitivity analysis is likewise performed. It is observed that outcomes do not change under various settling coefficient values, demonstrating that the proposed structure for area choice is very robust. The suggested methodology is verified and compared with results of other MCDM tools of prominence such as fuzzy TOPSIS and fuzzy VIKOR.

The criteria weights, dependent on experts' assessments, are calculated by fuzzy AHP. The proposed hybrid fuzzy TOPSIS–GRA technique is utilized to assess the location alternatives. A direct implication of this study is that this system might be utilized by various retail organizations in both the private sector as a strategy to expand their current offerings and chain of stores. Apart from the criteria considered in the study, the retailers might decide to add or remove any existing criteria in relation to their convenience store and unique user

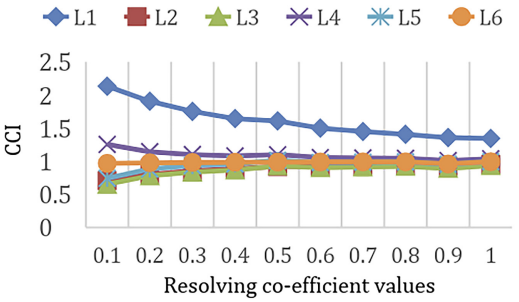


Figure 3.
Sensitivity analysis of CC_i values for each alternative

behaviour. The proposed methodology might be utilized in various metropolitan urban areas in India and abroad. The framework can be utilized by both the central and state government agencies to establish and expand their set of public ration stores.

There are some limitations to this research. Firstly, we have taken into account only eight selection criteria, but additional criteria such as environmental considerations, labour availability could further enhance ranking results. We have also considered experts' opinions based on the questionnaire and interviews. Secondly, the proposed method can be also be used and applied to evaluate location alternatives in other cities in India and also in other major cities of the world for establishing retail stores.

In future, different MCDM methods such as Multi-Objective Optimization on the basis of Ratio Analysis (MOORA), Elimination and Choice Translating Reality (ELECTRE), Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE), Complex Proportional Assessment (COPRAS) can be used as alternative performance evaluation techniques. Furthermore, Stochastic Multi-Criteria Decision-Making methods (SMCDM) such as Stochastic AHP, ANP and SVIKOR (Stochastic Visekriterijumska Optimizacija I Kompromisno Resenj) can be applied and compared to check the validity of our findings.

References

Arsovski, S., Todorovic, G., Lazić, Z., Arsovski, Z., Ljepava, N. and Aleksic, A. (2017), "Model for selection of the best location based on fuzzy AHP and Hurwitz methods", *Mathematical Problems in Engineering*, Vol. 2017, 2803461, p. 12.

Bai, C. and Satir, A. (2020), "Evaluating green supplier satisfaction", *Modern Supply Chain Research and Applications*, Vol. 2 No. 2, pp. 63-81.

Berman, B. and Evans, J.R. (1995), "Retail management: a strategic approach", Doctoral Dissertation, Univerza v Mariboru, Ekonomsko-Poslovna Fakulteta.

Buckley, J.J. (1985), "Fuzzy hierarchical analysis", *Fuzzy Sets and Systems*, Vol. 17 No. 3, pp. 233-247.

Burnaz, S. and Topcu, Y.I. (2006), "A multiple-criteria decision-making approach for the evaluation of retail location", *Journal of Multi-Criteria Decision Analysis*, Vol. 14, pp. 67-76.

Burillo, P. and Bustince, H. (1996), "Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets", *Fuzzy Sets and Systems*, Vol. 78, pp. 305-316.

Büyüközkan, G. (2004), "Multi-criteria decision making for e-marketplace selection", *Internet Research*, Vol. 14 No. 2, pp. 139-154.

Chakravorty, M. (2019), "Indian Retailers revisit consumer strategy from mass marketing to intelligence based targeted approach", *India Retail Forum, Deloitte India*, available at: https://www2.deloitte.com/in/en/pages/consumer-business/articles/retail-leadership-summit-2019.htm (accessed 26 August 2020).

Chang, H. and Hsieh, C.M. (2014), "A TOPSIS model for chain store location selection", *Review of Integrative Business and Economics Research*, Vol. 4 No. 1, pp. 410-416.

Chen (2019), "A new multi-criteria assessment model combining GRA techniques with intuitionistic fuzzy entropy-based TOPSIS method for sustainable building materials supplier selection", *Sustainability*, Vol. 11 No. 8, pp. 1-18.

Cottrell, J.L. (1973), "An environmental model for performance measurement in a chain of supermarket", *Journal of Retailing*, Vol. 49 No. 3, pp. 51-63.

Craig, C.S. (1984), "Models of the retail location process: a review", *Journal of Retailing*, Vol. 60 No. 1, pp. 5-36.

Dubois, D. and Prade, H. (1980), *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, Boston.

Erbıyık, H., Özcan, S. and Karaboğa, K. (2012), "Retail store location selection problem with multiple analytical hierarchy process of decision making an application in Turkey", *Procedia-Social and Behavioral Sciences*, Vol. 58 No. 1, pp. 1405-1414.

Ge, D., Hu, L., Jiang, B., Su, G. and Wu, X. (2019), "Intelligent site selection for bricks-and-mortar stores", *Modern Supply Chain Research and Applications*, Vol. 1 No. 1, pp. 88-102.

Gupta, S., Ahuja, G. and Kumar, G. (2018), "Identification of optimum locations for charging of electric vehicles", *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, 2018, pp. 867-872.

Gupta, S., Soni, U. and Kumar, G. (2019), "Green supplier selection using multi-criterion decision making under fuzzy environment: a case study in automotive industry", *Computers and Industrial Engineering*, Vol. 136, pp. 663-680.

Hoch, S.J., Kim, B.D., Montgomery, A.L. and Rossi, P.E. (1995), "Determinants of store-level price elasticity", *Journal of Marketing Research*, Vol. 32 No. 1, pp. 17-29.

Ingene, C.A. and Brown, J.R. (1987), "The structure of gasoline retailing", *Journal of Retailing*, Vol. 63 No. 4, p. 365.

Irwin, B. (1986), *Basic Retailing*, Irwin, Illinois.

Işıklar, G. and Büyüközkan, G. (2007), "Using a multi-criteria decision making approach to evaluate mobile phone alternatives", *Computer Standards and Interfaces*, Vol. 29 No. 2, pp. 265-274.

Jones, M., Mothersbaugh, D. and Beatty, S. (2003), "The effects of locational convenience on customer repurchase intentions across service types", *Journal of Services Marketing*, Vol. 17, pp. 701-712.

Karande, K. and Lombard, J.R. (2005), "Location strategies of broad-line retailers: an empirical investigation", *Journal of Business Research*, Vol. 58, pp. 687-695.

Kaufmann, A. and Gupta, M.M. (1991), *Introduction to Fuzzy Arithmetic: Theory and Applications*, VanNostrand Reinhold, New York.

Kaufmann, P.J., Donthu, N. and Brooks, C.M. (2000), "Multi-unit retail site selection processes: incorporating opening delays and unidentified competition", *Journal of Retailing*, Vol. 76 No. 1, pp. 113-127.

Kumar, A., Kashyap, H., Malhotra, H., Rawat, K., Kumar, G. and Soni, U. (2019), "Preferential selection of locations for installing CCTV using fuzzy TOPSIS approach: a case study for North Delhi", *2019 Twelfth International Conference on Contemporary Computing (IC3)*, Noida, pp. 1-7.

Kuo, R.J., Chi, S.C. and Kao, S.S. (2002), "A decision support system for selecting convenience store location through integration of fuzzy AHP and artificial neural network", *Computers in Industry*, Vol. 47 No. 2, pp. 199-214.

LBS (2013), "Hangout haven" available at: https://www.businesstoday.in/magazine/lbs-case-study/london-business-school-case-study-on-7-eleven/story/194769.html (accessed 15 November 2019).

Malik, S., Kumari, A. and Agrawal, S. (2015), "Selection of locations of collection centers for reverse logistics using GTMA", *Materials Today: Proceedings*, Vol. 2 Nos 4-5, pp. 2538-2547.

Mendes, A.B. and Themido, I.H. (2004), "Multi-outlet retail site location assessment", *International Transactions in Operational Research*, Vol. 11 No. 1, pp. 1-18.

Nyaoga, R., Magutu, P. and Wang, M. (2016), "Application of Grey-TOPSIS approach to evaluate value chain performance of tea processing chains", *Decision Science Letters*, Vol. 5 No. 3, pp. 431-446.

Olson, D.L. (2004), "Comparison of weight in TOPSIS models", *Mathematical and Computer Modelling*, Vol. 40, pp. 721-727.

Önüt, S., Efendigil, T. and Kara, S.S. (2010), "A combined fuzzy MCDM approach for selecting shopping center site: an example from Istanbul, Turkey", *Expert Systems with Applications*, Vol. 37 No. 3, pp. 1973-1980.

PWC (2017), "The promise of Indian retail", available at: https://www.pwc.in/assets/pdfs/publications/2017/the-promise-of-indian-retail-from-vision-to-execution.pdf (accessed 12 November 2019).

Roig-Tierno, N., Baviera-Puig, A., Buitrago-Vera, J. and Mas-Verdu, F. (2013), "The retail site location decision process using GIS and the analytical hierarchy process", *Applied Geography*, Vol. 40, pp. 191-198.

Shan, W., Yan, Q., Chen, C., Zhang, M., Yao, B. and Fu, X. (2019), "Optimization of competitive facility location for chain stores", *Annals of Operations Research*, Vol. 273 Nos 1-2, pp. 187-205.

Singh, R.K., Chaudhary, N. and Saxena, N. (2018), "Selection of warehouse location for a global supply chain: a case study", *IIMB Management Review*, Vol. 30 No. 4, pp. 343-356.

Torfi, F.A., Farahani, R.Z. and Rezapour, S. (2009), "Fuzzy AHP to determine the relative weights of evaluation criteria and Fuzzy TOPSIS to rank the alternatives", *Applied Soft Computing*, Vol. 10 No. 2, pp. 520-528.

Turhan, G., Akalın, M. and Zehir, C. (2013), "Literature review on selection criteria of store location based on performance measures", *Procedia - Social and Behavioral Sciences*, Vol. 99, pp. 391-402.

Tzeng, G.H., Teng, M., Chen, J.J. and Opricovic, S. (2002), "Multicriteria selection for a restaurant location in Taipei", *International Journal of Hospitality Management*, Vol. 21 No. 2, pp. 171-187.

Yang, C.L., Chuang, S.P., Huang, R.H. and Tai, C.C.. (2008), "Location selection based on AHP/ANP approach", *IEEE International Conference on Industrial Engineering and Engineering Management*, Singapore, pp. 1148-1153.

Yap, J.Y.L., Ho, C.C. and Ting, C.Y. (2019), "A systematic review of the applications of multi-criteria decision-making methods in site selection problems", *Built Environment Project and Asset Management*, Vol. 9 No. 4, pp. 548-563.

Corresponding author
Parnika Tyagi can be contacted at: parnikatyagi15@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

AIP Conference Proceedings

CVD synthesis and characterization of ultrathin MoS_2 film

Cite as: AIP Conference Proceedings **2276**, 020044 (2020); https://doi.org/10.1063/5.0026128
Published Online: 29 October 2020

Mohammad Imran, Nahid Chaudhary, Aurangzeb Khurram Hafiz, Bharti Singh, and Manika Khanuja

View Online Export Citation

ARTICLES YOU MAY BE INTERESTED IN

2-D graphitic carbon nitride nanostructures for optoelectronic application
AIP Conference Proceedings **2276**, 020006 (2020); https://doi.org/10.1063/5.0025687

Synthesis and characterization of Fe doped BCZT piezoceramic
AIP Conference Proceedings **2276**, 020011 (2020); https://doi.org/10.1063/5.0026153

Comparative study of photocatalytic activity of hydrothermally synthesized ultra-thin MoS_2 nanosheets with bulk MoS_2
AIP Conference Proceedings **2276**, 020030 (2020); https://doi.org/10.1063/5.0025756

Meet the Next Generation of Quantum Analyzers
And Join the Launch Event on November 17th

Register now
Zurich Instruments

AIP Publishing

© 2020 Author(s).

CVD Synthesis and Characterization of Ultrathin MoS_2 Film

Mohammad Imran[1], Nahid Chaudhary[1], Aurangzeb Khurram Hafiz[1], Bharti Singh[2] and Manika Khanuja[1, a)]

[1]*Centre for Nanoscience and Nanotechnology Jamia Millia Islamia (Central University), New Delhi- 110025, India*

[2]*Department of Applied Physics, Delhi Technological University, Delhi - 110042, India*

[a)]Corresponding author*: manikakhanuja@gmail.com*

Abstract: Transition Metal Dichalcogenides (TMDs) have been fascinating a wide range of scientific interest. Molybdenum disulfide (MoS_2) is the most focused TMDs. In this study, the ultrathin layer of MoS_2 was synthesized through sulfurization using molybdenum oxide (MoO_3) and sulfur(S) powder as precursors on SiO_2 substrate in self-assembled CVD at fixed temperature. The sample was further studied using various characterizations viz. Raman Spectroscopy, Scanning Electron Microscopy (SEM), Fourier Transform Infrared Spectroscopy technique (FTIR) and Photoluminescence (PL), which showed that MoS_2 can be fabricated on SiO_2 substrate easily by maintaining the temperature at 775˚C for 20 to 30 minutes.

INTRODUCTION

The structural arrangement of TMDs consists of one layer of transition metal atom sandwiched between two chalcogen atoms, where the layers are connected by the covalent bond and interlayers are attached by weak vander Waals forces, that leads to strong bonding within the layers and weak interlayer interactions [1-4]. MoS_2 is one of the TMD material that has been identified as nested fullerene-like nano-dots, 1-D nanotubes and can exists in three phases hexagonal, rhombohedral and trigonal. Usually, the hexagonal phase of MoS_2 is stable state, while rhombohedral and trigonal phases of MoS_2 are metastable state in nature [5-8]. MoS_2 has acquired considerable interest in the field of optoelectronics due to non-zero varying band gap. Bulk MoS_2 has an indirect band gap nearly ~1.3 eV while its deposited mono film possesses to show direct band gap of about 1.8 eV [3,9] In recent decade, two-dimensional (2D) nanosheet or only a single layer thick substance like boron nitride, graphene, silicene and TMDs have been fascinated in research areas due to their outstanding electrical properties. These materials may play as a key role in forthcoming technological advancement. Earlier, 2D materials were believed only as a theoretical and unstable in the material world. In 2004, Andre Geim and Kostya Novoselov et al. had observed, a flat honeycomb structured with hexagonal lattice, consisting of carbon atoms, named as graphene ($_6C^{12}$). Graphene was systematically studied and reported to have an interesting electrical property with zero band gap that will allow to withstand high current densities, follow Dirac equation for electron transportation and many other advantageous properties [10]. The zero bandgap of graphene is not suitable for electronic and optoelectronic applications. The fabrication of field effect transistor (FET) using graphene have reported to hold some critical issues during switching off process of transistors. In order to fabricate very efficient FET, a more suitable material is required [9,11-13].

Apart from graphene many other TMDs shows remarkable properties. The transition from indirect to direct bandgap of MoS_2 provides remarkable properties. Monolayered MoS_2 has been observed to show mobility nearly ~ 410 $cm2V^{-1}s^{-1}$ which is doubled as compared to bulk MoS_2 that shows mobility of ~ 200 $cm^2V^{-1}s^{-1}$ [1]. Several synthetic methods have been reported for the synthesis of mono and multilayer MoS_2 films using various approaches including chemical vapor deposition (CVD), thermolysis of $(NH_4)MoS_4$, electrostatic spinning, sulfurization of MoO_3, , physical vapor deposition (PVD), template method, hydrothermal method, micromechanical, intercalation and liquid phase exfoliation etc. [3,14-20]

In this paper, CVD method with atmospheric pressure was used to synthesize flower-like MoS_2 nanosheet using MoO_3 powder and sulphur powder as precursors. The reaction process and growth mechanism of MoS_2 nanosheets is optimized and discussed.

EXPERIMENTAL DETAILS

The growth of the MoS_2 layers were carried on on n-type (100) SiO_2 substrate inside a self-assembled CVD system.

International Conference on Advanced Materials
AIP Conf. Proc. 2276, 020044-1–020044-5; https://doi.org/10.1063/5.0026128
Published by AIP Publishing. 978-0-7354-4006-7/$30.00

A schematic diagram of experimental arrangement is as shown in fig. 1. The two SiO_2 substrates of 300 nm dense oxide (O_2) layer were utilized. Substrates were sonicated in acetone and ethanol for 5 minutes using digital ultrasonic cleaner, afterwards substrates were rinsed in DI and for further cleansing mild oxygen plasma treatment was done for 10 minutes. MoO_3 (0.25mg) powder (99.5%, sigma aldrich) was put inside a 2 centimeter deep ceramic boat, an n-type (100) Si substrate was kept on the same ceramic boat facing horizontally (H) downward, the other one was kept vertically upright at 90 degree with respect to tube of furnace and then positioned at the middle of the heater.

Another boat, holding 500 mg of sulphur powder (99.5%, sigma aldrich) was put in to a 4 inch diameter quartz tube and it was placed at the edge of hot coil (furnace) at a distance 12 cm away from the center of the furnace where MoO_3 powder source was placed.

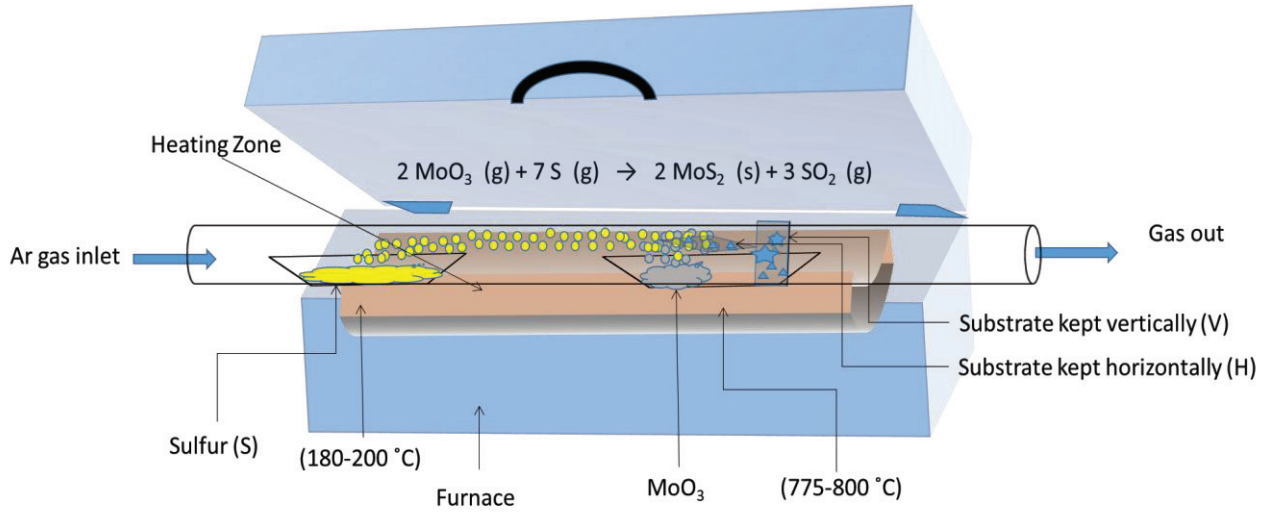


FIGURE 1: Schematic representation of the experimental setup for MoS_2 growth.

The furnace was heated up to 300 ˚C in 1 hour and up to 600 ˚C in another hour. After that it was raised to the step point value 775˚C in another 30 minutes and then it was held at that temperature for 20 minutes. After this, lid of the furnace set-up was opened to decrease the temperature within few minutes; the whole process was done in argon gas atmosphere which is kept at 10 sccm pressure.

CHARACTERIZATION

The morphological analysis of MoS_2 was carried out by the images using optical microscope (Lieca Microsystems CMS - DM4000 M LED) and scanning electron microscopy (SEM Zeiss, Sigma). Structural properties and bond formation were studied by FTIR (Vertex 70 V, Bruker, Optik) spectroscopy. PL and Raman spectroscopy techniques were carried out using Raman microscope (HR800 JY, Lab RAM HR) to see the difference in the peaks of MoS_2.

RESULTS AND DISCUSSIONS

Fig 2 (a, b) shows the optical microscope images of the as- synthesized MoS_2 nanosheets, which confirm the formation of MoS_2 nanosheets on n-type (100) silicon substrate. The optical images show the differences in the horizontal and vertical positioned of substrate. Further, the difference can be well observed by the FESEM images as shown in fig. 2 (c, d- horizontal alignment and e, f- vertical alignment). The SEM images shows flower like nanosheets morphology of the thin film of MoS_2. In case of horizontal alignment of substrate, thick nanosheets are observed. However, as shown in fig. 2(f), thin nanosheets are observed for vertical aligned substrate.

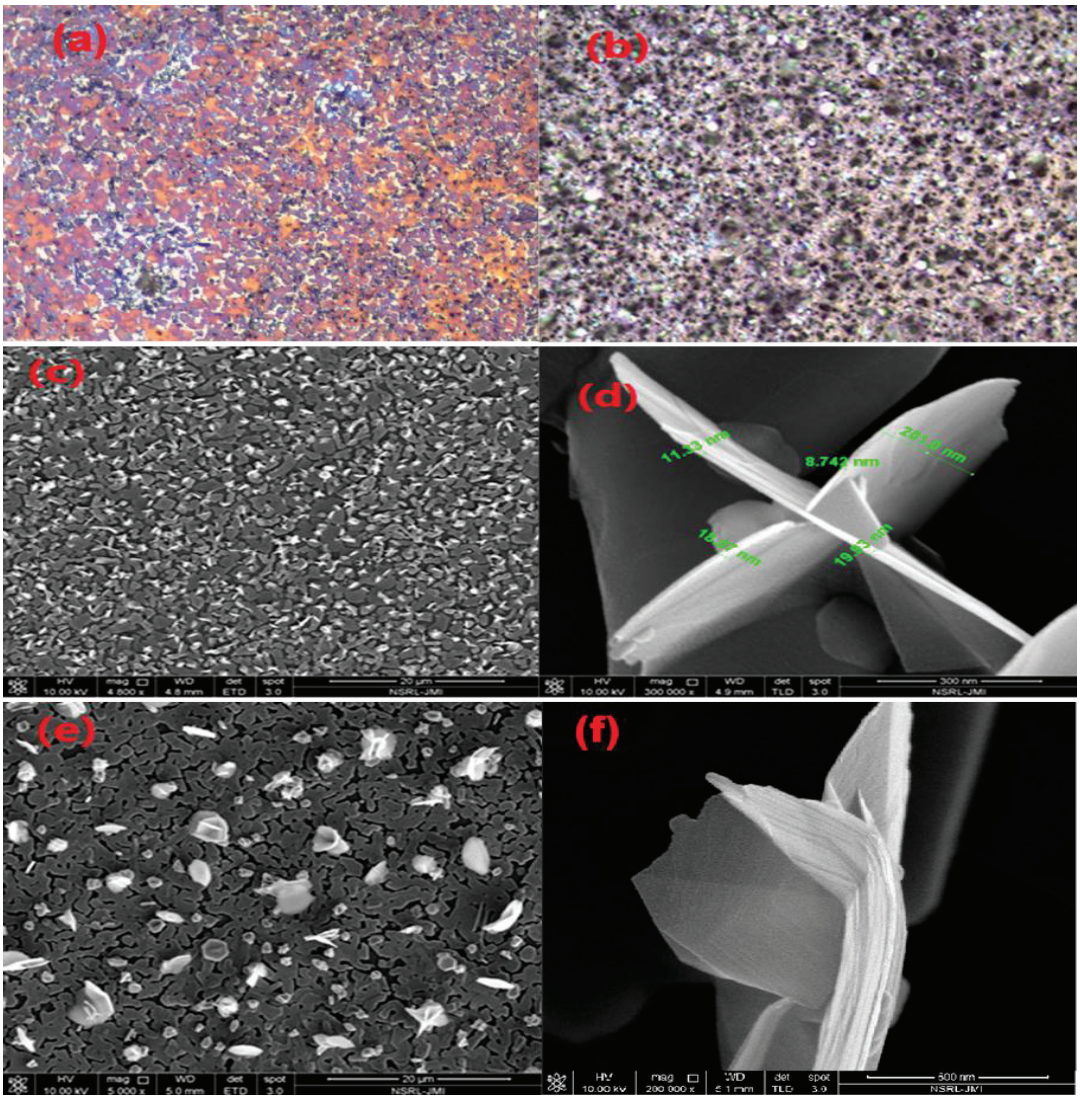


FIGURE 2: (a, b) Optical micrographs at 100x, (c, d) SEM images taken of 1st sample kept horizontally and (e, f) SEM images of 2nd sample taken vertically.

Fig.3 (a) show the FTIR transmission spectra, scanned in the spectral region from 400 cm^{-1} to 2000 cm^{-1}. The recorded spectra of MoS_2 thin film for vertical alignment positioned substrate shows absorption peak at 471 cm^{-1} and 518 cm^{-1} as shown in figure 3 (a) corresponding to the Mo-S bond present in the MoS_2 thin film and peak at 930 cm^{-1} are arises due to S-S bond in MoS_2. The stretching vibrations of Mo-O bond and the hydroxyl group attributed to the absorption bands at 1102 cm^{-1}, 1258 cm^{-1} and 1699 cm^{-1}.

Fig. 3 (b, c, d) show the Raman spectroscopic result of the MoS_2 thin films. The crystals growth for a vertically kept substrate shows Raman signature peaks at 385.46 and 409.66 cm^{-1} that corresponds to E^1_{2g} and A_{1g}, respectively while the peaks are slightly shifted right to 385.22 and 410.24 cm^{-1} for horizontally aligned substrate. The peak frequency difference between E^1_{2g} and A_{1g} for vertical and horizontal aligned sample is found to be 24.2 cm^{-1} and 25.02 cm^{-1}, respectively which shows the growth of ultra-thin MoS_2 layers on silicon substrate.

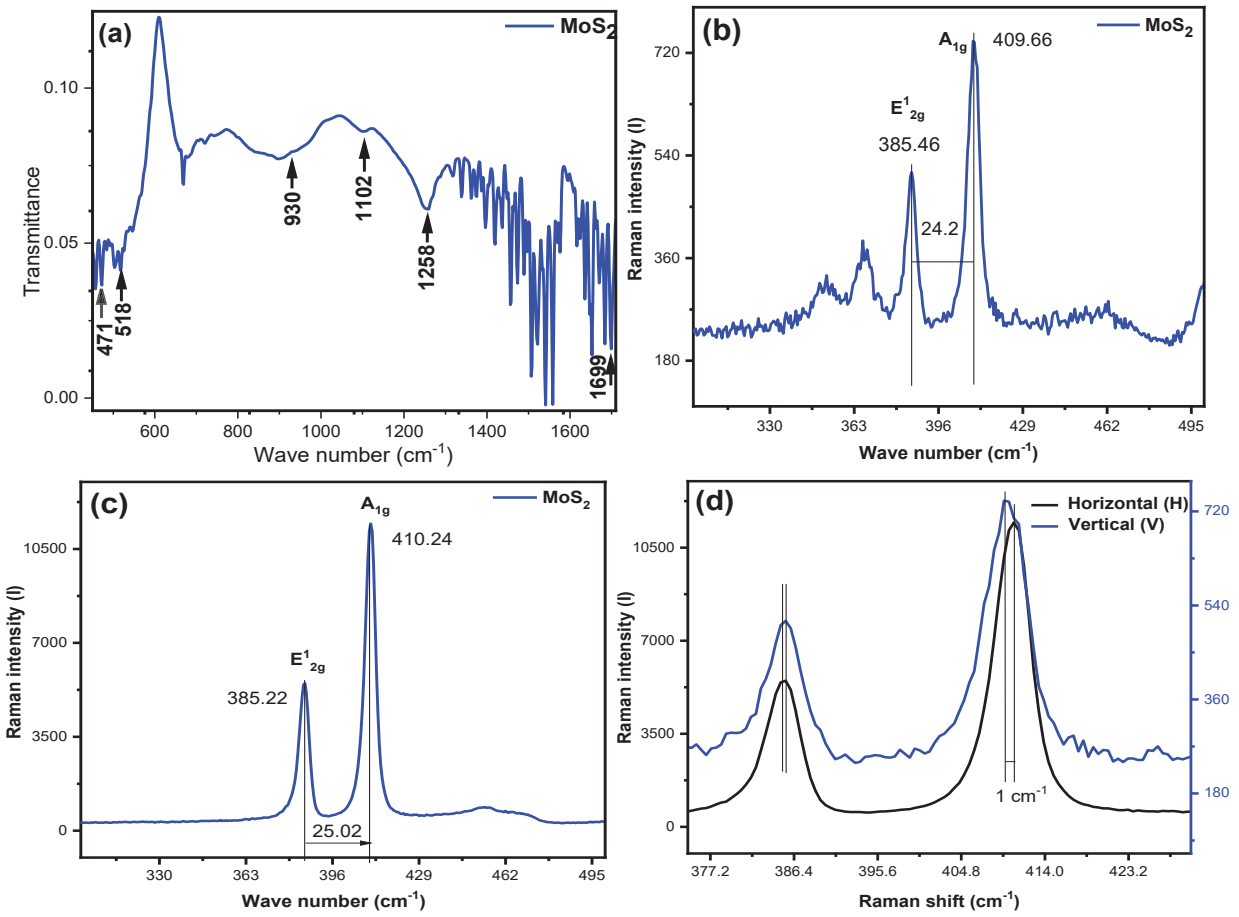


FIGURE 3: (a) FTIR transmission spectra and (b, c, d) Raman spectroscopic of MoS_2 thin film

CONCLUSION

In this paper, we have demonstrated the synthesis of ultrathin MoS_2 grown on silicon substrate by Chemical Vapour Deposition technique using MoO_3 and sulphur as precursor. The study of orientation of substrate in the CVD furnace during the synthesis process produces the morphological changes in the growth of the MoS_2 nanosheets is well studied.

ACKNOWLEDGEMENT

The present work is supported by University Grant Commission, New Delhi, India, project (No.F.4-5(201 FRP)/2015(BSR) and Science and Engineering Research Board (SERB) New Delhi, India, project (No. ECR/2017/001222).

REFERENCES

1. Wu, Wei, Debtanu De, Su-Chi Chang, Yanan Wang, Haibing Peng, Jiming Bao, and Shin-Shem Pei. Appl. Phys. Lett. 102, 142106, (2013)
2. Chhowalla, M.; Amaratunga, G. A. Nature 407, 164-167, (2000)
3. Nahid Chaudhary, Manika Khanuja∗, Abid, S.S. Islam, Sensors and Actuators A 277 , 190–198, (2018)
4. M. Nath, A. Govindaraj, C. N. R. Rao, Adv. Mater. 13, 283-286 , (2001)
5. Y. Golan et al., Adv. Mater. 11, 934-937 , (1999)
6. J. Xiao, D. Choi, L. Cosimbescu, P. Koech, J. Liu, and J. P. Lemmon, Chem. Mater. 22, 4522-4524, (2010)
7. J. Chen, N. Kuriyama, H. Yuan, H. T. Takeshita, T. Sakai, J. Am. Chem. Soc. 123, 11813-11814, (2001)
8. R. Tenne, L. Margulis, M. Genut, G. Hodes, Nature 360, 444 – 446 , (1992)

9. Mittal H, Kumar A,Khanuja M, Journal of Saudi Chemical Society, 23, 836-845 (2019), https://doi.org/10.1016/j.jscs.2019.02.004
10. Geim, A. K., and K. S. Novoselov. nature materials 6, 183-191, (2007)
11. K Chang, W. Chen, ACS Nano 5, 4720-4728, (2011)
12. Siddiqui I, Mittal H, Kohli VK, Gautam P, Ali M, Khanuja M, Mater. Res. Express. 5 (12), 125020, (2018)
13. Chaudhari N, Khanuja M, Islam SS, Polymer, 165, 168-73, (2019)
14. Ashraf, W., Fatima, T., Srivastava, K. and Khanuja, M., Applied Nanoscience, 9 (7), 1515-1529, (2019). DOI:10.1007/s13204-019-00951-4
15. L. Margulis, G. Salitra, M. Talianker, R. Tenne, Nature 365, 113 – 114, (1993)
16. M. Sun, J. Adjaye, A. E. Nelson, Appl. Catal. A 263, 131–143 , (2004)
17. Yang Song, Yingzi Peng, Suping You, Kewei Sun, Ji Chen, and Zhenghong Qian, Aip Advances, 5, 067119 (2015)
18. M. Remskar, A. Mrzel, Z. Skraba, A. Jesih, M. Ceh, J. Demsÿar, P. Stadelmann, F. Levy, D. Mihailovic, Science 292, 479 –481 , (2001)
19. A. N. Enyashin, S. Gemming, M. Bar-Sadan, R. Popovitz-Biro, S. Y. Hong, Y. Prior, R. Tenne, G. Seifert, Angew. Chem. Int. Ed.46, 623 –627, (2007)
20. Gonglan Ye, Yongji Gong, Junhao Lin, Bo Li, Yongmin He, Sokrates T. Pantelides, Wu Zhou, Robert Vajtai and Pulickel M.Ajayan, Nano Lett. 16, 1097−1103, (2016)

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Depression and Impaired Mental Health Analysis from Social Media Platforms using Predictive Modelling Techniques

Vaibhav Jain
Department of Information Technology
Bhagwan Parshuram Institute of Technology
Delhi, India
vaibhavces@gmail.com

Dhruv Chandel
Department of Information Technology
Delhi Technological University
Delhi, India
dhruvchandelon5@gmail.com

Piyush Garg
Department of Information Technology
Bhagwan Parshuram Institute of Technology
Delhi, India
gargpiyush03@gmail.com

Dinesh Kumar Vishwakarma
Department of Information Technology
Delhi Technological University
Delhi, India
dinesh@dtu.ac.in

***Abstract*— Depression is the leading global disability, and unipolar (as opposed to bipolar) depression is the 10th leading cause of early death, as stated by the World Health Organization (WHO) in 2015. The study aims to build an approach for depression and impaired mental health analysis from social media platforms. Although for Depression analysis and cure. Psyscologists preferred over machines because they are manipulative and precautionary to Human emotions to a greater extent, Machine Learning has an added advantage. It has no emotions; it studies patterns, not face or beauty or other factors. It studies a wide variety of data and then trains to give better predictions. Although it is not 100% reliable nor are the doctors. Moreover, in countries like India where people don't treat Depression as a Chronic Illness or don't even consider it as an illness of any sort, embedding Machine Learning Depression Detection Algorithms in Social Media combined with recommendation systems to treat a Human Mind positively, still being unnoticeable is a Great Boon to humanity The study is assisted by data collected from users after obtaining their consent and applying data preprocessing techniques. Several machine learning is used to analyze the data in the best way possible. A VAPID Technique is developed that performs far better than a classic feed-forward neural network. This study aims to develop a correlation between features and depressed people to observe a continuous pattern. Moreover, the aim is to conclude that social media can be a new exceptional methodology for analyzing depression and analyzing indirect patterns, improving many lives.**

Keywords—Depression, Mental Health, Machine Learning, Data Preprocessing

I. INTRODUCTION

Depression is a very severe and unattractable illness that is spreading rampantly, especially among this generation. It was pragmatically found out that 264 million people of all ages suffer from depression (World Health Organization, 2017). [1] .[2] Approximately 76% to 85% of depressed people in low and middle-income countries are not receiving treatments; [2] barriers to treatment include inaccurate assessment, lack of trained health care providers, social stigma, and lack of resources. [3] Rates of mental health workers vary from below 2 per 100,000 populations in low-income countries to over 70 per 100,000 in high-income countries. Even after so many advancements, there is still no full-proof method to study different mental illnesses; It is currently based highly on patients' self-provided reports or by people in their vicinity.

[1] Depression is ranked as the single largest contributor to global disability (7.5% of all years lived with disability in 2015). At its worst, depression can lead to suicide; over 800,000 people die due to suicide every year. Suicide is the second leading cause of death in 15-29-year-olds. In India, the National Mental Health Survey 2015-16 reveals that nearly 15% of Indian adults need active intervention for one or more mental health issues and one in 20 Indians suffers from depression. Estimation presents that in 2012, India had over 258,000 suicides, with the age-group of 15-49 years being most affected.

Social media users have passed the 3.8 billion mark. Therefore, social media's perspective is being explored to find out possible manifestations of depression in individuals. People use social media to express themselves, including their thoughts, opinions, mood, and daily lifestyle; this information can be used to know about their psychological condition and even create a possible method for detecting mental disorders in the earlier stages of development. This can be done by monitoring the nuances that occur in their daily posts, including their facial expressions, background, preference of color, caption, hashtags used, emoji counts, and many more. All this information can be evaluated together and condensed to form a mathematical model that can be applied to detect and even verify their mental condition. The main contributions of the work are as follows:

- Create a structural methodology which can detect depression from data collected via social media platforms of users
- Proposing some expected results or patterns which can be concluded from depressed peoples.
- Finding the best fit predictive model which can analyze people's dataset and can predict whether they are depressed or not

A. Parameters For Selection Of Best Platform

Platform choosing to get the best result is very important because different social media platforms have different system designs, and because of that, every platform asks for different information from their users, so choosing a social media platform was an issue in itself. e.g., [2] have used twitter[3] to conduct similar research; one thing necessary is that twitter has a vast database, and it is easy to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

find useful data on twitter. However, alternative features are needed to make the results further precise as only textual data is not enough for conducting research, so for that purpose, an ideal platform is needed with more visual content and more demographically explaining features. Some points for using Instagram can be:

- Instagram demographics have increased exponentially over any other available platform; it will be the world's most widely used social media platform.
- Unlike any other platform, it has multiple features like images, videos, texts, hobbies, and many more. So, depression analysis will be more precise.
- Also, people have more trust on Instagram than any other social media platform about their social security. <add the reference here> Therefore, people feel more comfortable to express their true feelings on Instagram than on twitter.
- Instagram's auto-captioning backend model creates a powerful expressible statement that generates text relevant to objects and faces in images, which can be an asset to the paper.

II. RELATED WORK

Several papers and journals have published various documents related to depression analysis using social media using several different models, and this topic has always remained a debatable topic. Reece et al. [4] has analyzed depression over a dataset obtained using amazon Mechanical Turk, in which they used discrete image features obtained from user's profiles and some other data related to Instagram such as followers, following, likes, and comments, and applied semi-supervised model over that dataset in which they obtained their best results from a 100. Tree sized Random Forest in which they attained the best accuracy of around 70%. Bayesian logistic regression with uninformative priors was also used to determine the strength of two individual predictors. Kleftaras et al. [5] surveyed the people of Greece, where they used a questionnaire created by PichoAtEl(1984) [6] in 2 languages, Greek and English, by the help of 2 psychiatrists who also rated their reviews about people. [5] It consists of 52 items that describe various depressive symptoms: affective, cognitive, and somatic (e.g., "I have difficulties in avoiding the negative thoughts that come to my mind"). High scores relate to more severe levels of depression. Later on, they conducted Principal Component analysis and multiple regression on the dataset, and they concluded their results as: -

- The higher the problematic SNS usage, the higher the depressive symptomatology
- Based on conducted multiple regression, they concluded that Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness, contributes 33% depressive symptomatology and also that residence and gender is a purely independent feature for depression.

Morrison et al. [7] have also portrayed their results upon excessive internet usage(especially social media) and depression, specifically among the adolescents, to find out they circulated a three scaled online questionnaire consisting of the IA Test, the Internet Function Questionnaire [8], and the Beck Depression Inventory (BDI) in which 1,319 respondents completed the questionnaires, with 18 (1.2%) identified as falling in the Internet Addictive category. They conducted a factorial test between IA and NA (Non-addicted audience). They concluded that men are prone to IA among adolescents, which causes depression and the IA group engaged significantly more than the NA group in sexually gratifying websites, gaming websites, and online community/chat websites. Moreover, many other authors have published their different views over similar topics such as [9] Tao et al. have targeted the internet media platforms which have textual content also as features and completed sentiment analysis or [3] who have deeply analyzed their audience's mood oscillation while in depression or either it is .O'Keeffe GS et al.[10] who improvised their study by expanding their audience to individuals, their family, living environments, and Ryu EJ et al. [11]has also done work on displaying how depression pushes people's mental state so bad to commit suicide.

III. DATA

In data analyses, one of the fundamental pillars is data. Model accuracies and hypothesis depends very much on how the data looks like in term of statistical terms. This section will cover all points that how a state-of-the-art dataset is made, what were that preprocessing methodology and technique applied to attain the best results

A. Data Collection

Data Collection is one of the essential and challenging steps because participants are generally afraid of data theft. Every user was asked for their consent, maintaining agreement that they will solely have the right to their data, and there will be no sharing of this data to anyone in any case except publishers of this paper—documentation of 159 users' data done during the whole process.

The users are from Indian origin; hence most of them are bilingual in which "Hindi" is the mother tongue of many, so post captions are translated and bio into English if the text was in any language except English using google translate API.[12]

B. Data Preprocessing

Raw data extracted is usually not an ideal dataset for researching since it is full of missing and inaccurate values for which an in-depth cleaning process is required.

1. Attaining required batch size: - For feeding the dataset into an LSTM, the dataset is needed to make a 12 batch sized data so for that purpose, users with a post count of less than 9 (at least 70% of the required batch size) were rejected but users who were having posts greater than ceil value, their data was duplicated with an average of values so far.
2. Emoji Count –Willoughby et al. [13] suggest that text context decides emoji's value. At some business conversations, emojis do not play a vital role, but peoples in relationships use different emojis, and at that stage, emojis can have very vast and different meanings. Also, the sarcastic text is an issue at that moment. Emojis plays up a very significant role [13]. Thompson and Flik et al. [14] Huang et al. [15] proved that fellows or co-workers in enjoyment frequently use emoji, and that result matched with the dataset as data

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

consist mostly of fellows of college students. [7] Morrison et al. have done an in-depth study by making CNN predictive model on the Facebook dataset, concluding that emojis played a profound role in predicting emotions of people happy and said from which can be concluded people depressed or not, also [3] Giota KG et al. have resulted in similar results.

1. Translation - For leveraging bilingual data, sentimental analysis requires a common language. Although most users were Hindi speakers, still preferred to convert all of the textual data to English because an analysis proved that 77% of all the textual data was already in the English language. So for that, google translate API[19] is used to translate the data.
2. Sentiment Analysis- Used [12], which on feeding text labels the data positive or negative, depicting that either the text is having a positive or negative emotion

C. *Visual Features*

Data visuals played a significant role in the dataset, the prime reason for choosing Instagram, as discussed in section *I-A* data. The data was analyzed, and seventeen features were described, displayed in Table-1. Furthermore, the problem was finding out which features should be considered for finding out about depressed people; Benjamin et al.[16] shows that high hue, low brightness, and saturation are common patterns observed in depressed people. Color quality metrics and the HVS parameters can display about psychophysical data, and referring several other papers, a result is reached on and decided to these features: -

TABLE-1-DEPICTING ALL VISUAL FEATURES

S.NO.	FEATURES
1.	Q1_RED_CHANNEL
2.	MEDIAN_RED_CHANNEL
3.	Q3_RED_CHANNEL
4.	Q1_GREEN_CHANNEL
5.	MEDIAN_ GREEN _CHANNEL
6.	Q3_ GREEN _CHANNEL
7.	Q1_BLUE_CHANNEL
8.	MEDIAN_ BLUE _CHANNEL
9.	Q3_ BLUE _CHANNEL
10.	Q1_GRAYSCALE
11.	MEDIAN_ GRAYSCALE
12.	Q3_ GRAYSCALE
13.	HLS_HUE
14.	HLS_LIGHTNESS
15.	HLS_SATURATION
16.	HSV_HUE
17.	HSV_VALUE
18.	HSV_SATURATION

1. Channel Intensity of each color of any image: - Every Image consists of many pixels where for every pixel, having three colors Red, Green, Blue; by normalizing the average of each color separately, nine values are extracted.
2. HSL (Hue, Saturation, Lightness)- HSL transform of any image into 3-dimensional values, each specifying hue, saturation, and lightness [17]
3. HSV (Hue, Saturation, Value) -HSV values depict us as the method in which distinct colors mix, *saturation* dimension corresponding to the varied tints of glowingly colored paint, and the *value* dimension is resembling the blend of those paints with varying quantity of black or white paint.

Another visual description that can play a significant role is what the image is trying to describe. For example, the number of peoples in the image describes how gregarious a person is, which is also an essential factor of consideration. <, so, for that information, image parameters were unable. iska mtlb explains krdo ajeeb si lag ri hai padne mei> Some technique is needed to evolve which can give us an accurate description of the image instead of what the user has captioned it. For that cause, referred to previous works, Jain et al.[18] has evolved a unique method using skin tone and facial detection technique, which on low luma light and high luma skin tones work amazingly, or [19] have used neural net bordering and sensitivity technique for finding faces up to 90% accurate. However, Instagram's backend server has a potent pre-trained captioning model, so it is preferred to use it instead of creating a unique model.

Below Fig.1 and Fig.2 show very significant results from the database. These plots were constructed by converting each image collected from the user to the same dimension image and then taking the average of RGB values of each pixel; and in such ways, figures conclude that blue color has significantly more density then rest of the colors and equal parts of RGB makes dark black color. [20] Furthermore, black color signifies depression, darkness, seriousness, conventionality, unity, sorrow, mourning in the west, whereas blue color signifies for non-threatening color, which has a tranquillizing effect on the mind. It symbolizes depression, tranquillity, trust, confidence, conservatism, and dependability [20].

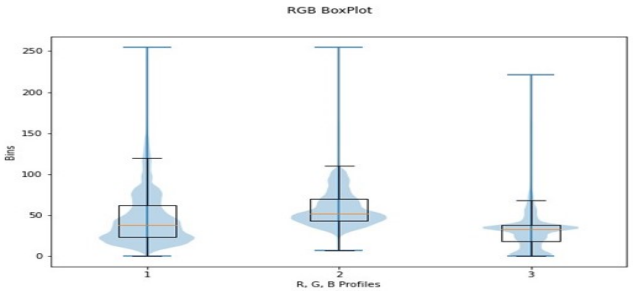


Fig-1-Depicting RGB box plot of the sample image

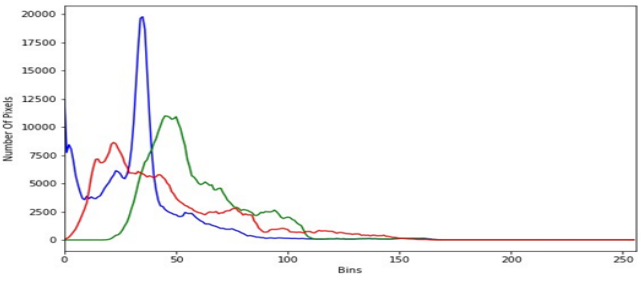


Fig-2-Depicting RGB Histogram for sample Image

IV. IMPLEMENTATION DETAILS AND PROPOSED METHODOLOGY

Many different approaches and flow of controls were implemented to achieve the best results and predictions but

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

the model that was finally rested upon was a Multimodal approach whose flow diagram us shown in Fig-4.

A. *Model 1*

Initially 'Model-1' was proposed as Logistic Regression Model. It was used to predict overthinking frequency. Logistic Regression worked to find linearity with its features and performs best with independent features, but the model had several features which were connected to each and for predicting could have needed to account all of them; also the dataset is very prone to overfitting because there were many outliers in the comparatively small-sized dataset that even after heavy feature engineering failed to improve performance of the model using Logistic Regression. In that case, a better model was required to get the best results.

To better cater Dataset outliers and problems related to the small dataset overthinking count is classified in 5 major categories as shown in Fig 3. Now Random Forest [21] model was implemented over the dataset. Random Forest is an ensemble tree technique that operates by many decision trees and gives results by mode for classification and mean or average for regressive results. Random Forest reduces training overfitting and training time; hence was preferable to use. Multiple small decision trees are combined; they contain accurate and inaccurate samples, trees protect each other from individual errors and produce precise results regarding other tree algorithms. Using hyperparameters tuning of various tree depth and leaf nodes, concluded that using entropy criterion as a measure of disorder and have min samples leaf as two can attain 65% accurate result on the testing dataset and can observe the confusion matrix image which is describing how many times peoples overthinking per week.

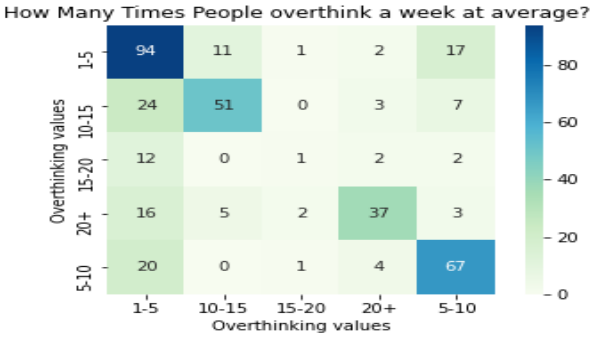


Fig-3-Heatmap of Confusion Matrix obtained from random forest classifier

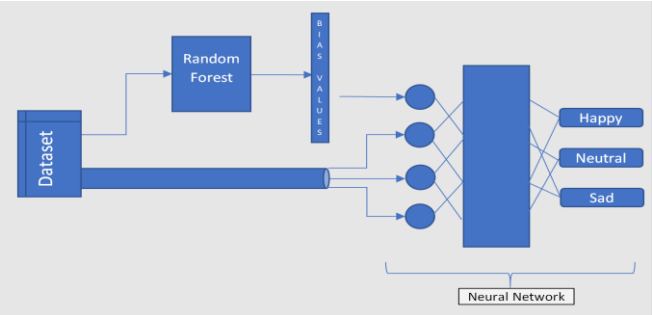


Fig. 4-depicting flow diagram of complete model architecture

TABLE-2-RESULTS OBTAINED FROM RANDOM FOREST CLASSIFIER

S.NO	CAT	P	R	F1-SCORE	SUPPORT
1.	**1-5**	0.57	0.75	0.65	125
2.	**10-15**	0.76	0.60	0.67	85
3.	**15-20**	0.20	0.06	0.09	17
4.	**20+**	0.77	0.59	0.67	63
5.	**5-10**	0.70	0.73	0.71	92
6.	**AC**	0.60	0.54	0.65	382
7.	**MA**	0.60	0.55	0.56	382
8.	**WA**	0.66	0.65	0.65	382

Abbreviations – CAT: Categories, P: Precision, R: Recall, AC: Accuracy, MA: Macro Average, WA: Weighted Average

B. *'Model-2'*

Initially, 'Model-2' was proposed as a Linear Regression Model but the dataset constraints didn't provide an adequate pattern to linear regression Model and trained model every time was either underfit or overfit. The second model tryout was LSTM based regressive model. It used a time-series feature of the dataset (the fact that only 12 most recent posts from every user and done duplication if the user has posted less than 12) and used depression score as the target value. It was observed that people in higher depression states were posting at extremities, i.e. either too often or rarely post. This lead to LSTM timestamp-based prediction produced inaccurate predictions. Moreover, many time-independent features in the dataset were colliding with time-based features and decreased the overall accuracy, so preferred to discard LSTM based approach.

Finally, the feed-forward Neural Network Model was taken into consideration. Regression-Based Neural Network was tested, but to obtain accurate predictions improvised it to a classification model by classifying data using the standard depression classifiers rather than numerical rating.

A number of layers were chosen by brute-forcing Neural Network through a series of possible layer sizes and depths. The given architecture was a sweet spot for training and testing as any further changes in layer deployment reduced Models' performance significantly by underfitting or overfitting the Model every time. Deep Neural Networks were not implemented to incur issue of Small dataset and prevent overfitting. Attaching a Softmax layer in the end layer and also the results of the "Model-1" which has five categorical data were feed in this model, this step was essentially done because very high data redundancy is observed in the model due to high variation in the dataset so "Model-2" was highly lagging of any base value over which it can conclude its result. So Random forest's result on a particular Instagram Post acted as the base value for Neural Network which helped us in attaining better overall performance of the Model. The final results have training accuracy of 69.46% and 59.55%. Table 2.1 has shown specification used with the model

TABLE-3-FEATURES AND SPECIFICATIONS OF THE NEURAL NETWORK

S.NO.	FEATURES	SPECIFICATION
1.	Loss Type	Categorical Cross-Entropy
2.	Epochs	50
3.	Optimizer	Adam
4.	Batch Size	32
5.	Data Shuffling	True

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

6.	Total Layers	9
7.	Dropout Layers	3
8.	Dense Layers	6

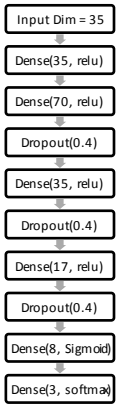


Fig-5-Layerwise Architecture of Neural Network with Activation Functions Alongside each layer

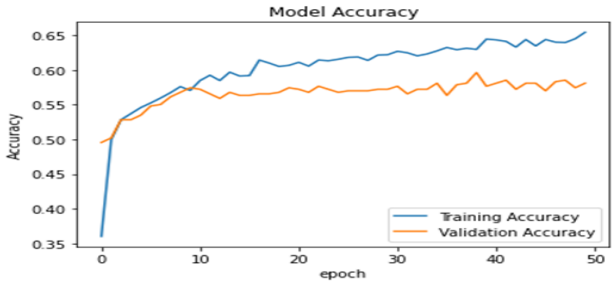


Fig-6-Model accuracy of Neural Network

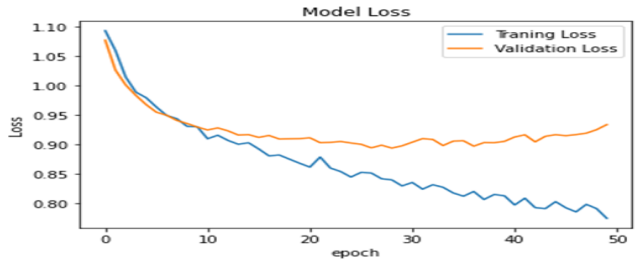


Fig-7-Model loss of Neural Network

V. Training And Testing Details

A. Model Training Approach

In predictive modeling, breaking down a dataset is one of the most critical steps because how one segregate their dataset decides randomization over a dataset and its precision

The data consists of 1908 entries, 12 entries for each user providing 159 peoples in the dataset, and 42 features for each entry. Out of those 42 features, 17 features are image-based features, which are described in later sections briefly now to maintain of perfect ratios of depressed and non-depressed people in both training and testing set; some of the undersigned methodologies have been followed: -

- The Dataset was randomly shuffled and Training: Testing Sub-Datasets were divided in the ratio of 8:2.
- The second approach used 10 percent of the training dataset to find the model's validation accuracy using the k-fold cross-validation algorithm.
- Thirdly some of more profiles are scraped randomly from Instagram using #depression and related hashtags and tested the model accuracies on those datasets.

B. Sample Workflow

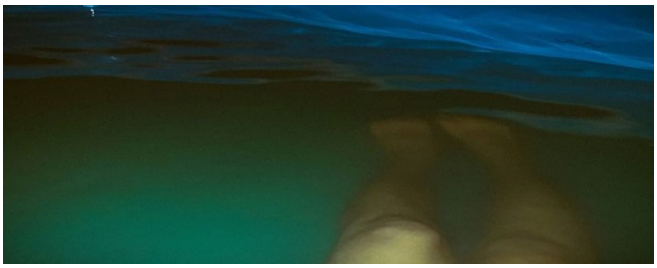


Fig-8- Sample Image of one of the users

The Image in 'Fig-8' is taken from the account post of an undisclosed user suffering from depression to better understand the workflow of the Model.

Image Caption: -Stasis in darkness/ Then the substanceless blue/ Pour of tor and distances.

TABLE 4-SENTIMENT ANALYSIS OVER THE CAPTION (PROBABILITY FOR EACH SENTIMENT)

S.NO.	NEGATIVE SENTIMENT	NEUTRAL SENTIMENT	POSITIVE SENTIMENT	LABEL
1.	0.660930818	0.20787287	0.339069182	neg

TABLE 5-CAPTION'S EMOJI USED AND LENGTH ANALYSIS (COUNT OF EMOJIS IN EACH TEXT)

S.NO	HAPPY EMOJI	SAD EMOJI	ROMANTIC EMOJI	ANGER EMOJI	LENGTH OF CAPTION
1.	0	0	0	0	235

TABLE-6-EXTRACTED DEFAULT ALTERNATE CAPTION OF THE IMAGE

S.NO	AC	HC	CT	TA	#C
1.	Photo by ********* on June 19, 2020. The image may contain: one or more people, water, and outdoor	1	0	0	0

Abbreviations - AC: Alternate Caption, HC: Human Count, CT: Contains Text (Boolean), TA: Tagged Accounts, #C: Hashtag Count

TABLE 7- POST'S BASIC INFORMATION

S.NO	TIMESTAMPS	LIKES	COMMENTS
1.	1.59E+09	78	0

TABLE-8-USER'S ACCOUNT INFORMATION

S.NO	USER NAMES:	BIO	FWS.	FWG.	#P
1.	******	This is an aesthetic dump yard of things I love.	387	297	237

Abbreviations: FWS: Followers, FWG: Following, #P: Post Count

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

VI. CONCLUSION

Out of many rising social media platforms, this research explored techniques to predict users' mental state based upon their semantic responses images, captions, bilingual behavior usage, and several other features. Statistical analysis over data concluded certain results:

- Precise conclusion of the user's mental state from a combination of its text and visual features; only semantic analysis is not enough for concluding mental states.
- Generally, images with dark colors such as grey, black, or dark filters displayed people in depression.
- When there is data with high variation, one can use the newly introduced "VAPID TECHNIQUE" to combine a model's prediction value to create a high bass in neural net models.

Predictions made using the proposed Methodology was about 60% accurate and can be trained and tester further if better datasets are open-sourced for Research purposes.

Furthermore, analyzing responses of crowdsourced data displayed that people in depression generally discuss lack of sleep, tiredness, and cuss words as reasons and symptoms of depression [22]. The results' prime limitation is the lack of dataset due to discontinuation of Instagram API by Facebook after acquiring it. It was very tough to get dataset and conduct research as individual entities were required to be consented to use their Instagram profile information as a part of this research.

REFERENCES

[1] "10 facts on mental health." https://www.who.int/news-room/facts-in-pictures/detail/mental-health (accessed Sep. 08, 2020).

[2] W. Yang and L. Mu, "GIS analysis of depression among Twitter users," *Appl. Geogr.*, vol. 60, pp. 217–223, Jun. 2015, doi: 10.1016/j.apgeog.2014.10.016.

[3] "Twitter. It's what's happening / Twitter." https://twitter.com/ (accessed Sep. 08, 2020).

[4] A. G. Reece and C. M. Danforth, "Instagram photos reveal predictive markers of depression," *EPJ Data Sci.*, vol. 6, no. 1, p. 15, Dec. 2017, doi: 10.1140/epjds/s13688-017-0110-z.

[5] K. G. Giota and G. Kleftaras, "The role of personality and depression in problematic use of social networking sites in Greece," *Cyberpsychology*, vol. 7, no. 3, 2013, doi: 10.5817/CP2013-3-6.

[6] P. PICHOT, P. BOYER, C. B. PULL, W. REIN, M. SIMON, and A. THIBAULT, "Un questionnaire d'auto-évaluation de la symptomatologie dépressive, le questionnaire QD2. II: Forme abrégée QD2A," *Rev. Psychol. appliquée*, vol. 34, no. 4, 1984.

[7] C. M. Morrison and H. Gore, "The relationship between excessive internet use and depression: A questionnaire-based study of 1,319 young people and adults," *Psychopathology*, vol. 43, no. 2, pp. 121–126, Feb. 2010, doi: 10.1159/000277001.

[8] S. M. Li and T. M. Chung, "Internet function and Internet addictive behavior," *Comput. Human Behav.*, vol. 22, no. 6, pp. 1067–1071, Nov. 2006, doi: 10.1016/j.chb.2004.03.030.

[9] X. Tao, X. Zhou, J. Zhang, and J. Yong, "Sentiment Analysis for Depression Detection on Social Networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10086 LNAI, Springer Verlag, 2016, pp. 807–810.

[10] G. S. O'Keeffe *et al.*, "Clinical report - The impact of social media on children, adolescents, and families," *Pediatrics*, vol. 127, no. 4. Pediatrics, pp. 800–804, Apr. 2011, doi: 10.1542/peds.2011-0054.

[11] E. J. Ryu, K. S. Choi, J. S. Seo, and B. W. Nam, "The relationships of Internet addiction, depression, and suicidal ideation in adolescents.," *Taehan Kanho Hakhoe Chi*, vol. 34, no. 1, pp. 102–110, 2004, doi: 10.4040/jkan.2004.34.1.102.

[12] "Cloud Translation | Google Cloud." https://cloud.google.com/translate/ (accessed Sep. 08, 2020).

[13] J. F. Willoughby and S. Liu, "Do pictures help tell the story? An experimental test of narrative and emojis in a health text message intervention," *Comput. Human Behav.*, vol. 79, pp. 75–82, Feb. 2018, doi: 10.1016/j.chb.2017.10.031.

[14] D. Thompson and R. Filik, "Sarcasm in Written Communication: Emoticons are Efficient Markers of Intention," *J. Comput. Commun.*, vol. 21, no. 2, pp. 105–120, Mar. 2016, doi: 10.1111/jcc4.12156.

[15] H. A. H., Y. D. C., and Z. X., "Exploring the potential effects of emoticons," *Inf. Manag.*, vol. 45, pp. 466–473.

[16] S. B. Guessoum *et al.*, "Adolescent psychiatric disorders during the COVID-19 pandemic and lockdown," *Psychiatry Research*, vol. 291. Elsevier Ireland Ltd, Sep. 01, 2020, doi: 10.1016/j.psychres.2020.113264.

[17] C. Abidin, "#In$tagLam: Instagram as a Repository of Taste, a Burgeoning Marketplace, a War of Eyeballs," in *Mobile Media Making in an Age of Smartphones*, New York: Palgrave Macmillan US, 2014, pp. 119–128.

[18] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 696–706, May 2002, doi: 10.1109/34.1000242.

[19] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 1, pp. 23–38, 1998, doi: 10.1109/34.655647.

[20] N. Singh and S. K. Srivastava, "Impact of Colors on the Psychology of Marketing — A Comprehensive over View," *Manag. Labour Stud.*, vol. 36, no. 2, pp. 199–209, May 2011, doi: 10.1177/0258042x1103600206.

[21] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, Jan. 2005, doi: 10.1080/01431160412331269698.

[22] O. Aydin, F. P. Çökmüş, K. Balikçi, D. Sücüllüoğlu-Dikici, and P. Ünal-Aydin, "The problematic use of social networking sites associates with elevated symptoms in patients with major depressive disorder," *Int. J. Soc. Psychiatry*, vol. 66, no. 5, pp. 496–503, Aug. 2020, doi: 10.1177/0020764020919791.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:51:02 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344693927

Design and Analysis of Rear Drive Axle of Go-Kart

Article · November 2020

CITATIONS
0

READS
14

1 author:

Shubham Singh
Delhi Technological University
1 PUBLICATION **0** CITATIONS

SEE PROFILE

All content following this page was uploaded by Shubham Singh on 16 October 2020.

The user has requested enhancement of the downloaded file.

Design and Analysis of Rear Drive Axle of Go-Kart

Shubham Singh
Department of Mechanical Engineering
Delhi Technological University, India

Abstract— Go-Kart is a light weight, single seater, low cost, four wheeler, and open cabin vehicle without any suspension and differential gear box. It is generally powered with single cylinder IC engine of maximum capacity of 135 CC. It is very simple in construction and equipped with semi floating, solid rear axle that subjected to the static and dynamic loading. This paper is mainly focus on the design of rear drive axle of the go kart and its various possible load conditions. The rear axle of go-kart is very crucial part because it is responsible for the kart motion and also bear a large fraction of total load. This study helps to design the rear solid axle of the go-kart for the safe functioning during operation and also a deep understanding about various loading scenarios.

Keywords: Go-Kart

I. INTRODUCTION

In Go-Kart competition, team of B-tech or diploma students design and manufacture their kart for participating in this championship. The mission of this championship is to encrypting students with technical as well as practical aspects of design and manufacturing a go-kart. This challenge is on global platform where the various national and international college teams are participate and show their best of the talent thought design and manufacturing on international level. Reason of this publication is to study about the various possible loading condition and its combined effect on axle during accelerating, climbing, turning and braking. So that to design our solid rear axle of go-kart. Here we prefer to use of solid shaft instead of hollow, because solid shaft has more torsional rigidity as compare to the hollow and hollow shaft is larger in diameter as compare to solid shaft for same strength, thus hollow shaft is required more bigger size of bearings and also required more space.

The following are the conditions of various loading on rear axle:-

- Bending moment due to weight shift of kart during accelerating, climbing and turning.
- Bending moment due to torque transmission through chain and sprocket.
- Bending moment due to resisting torque exerted by the caliper on disc brake.
- Twisting moment due to engine torque and braking
- Here we are using double bearings at each side of the shaft that increase the rigidity at some extent. And all the shear force and bending moment are transfer to the chassis through the bearings. Thus due to vertical load at the extreme ends of the axle (wheel hubs), the inside portion of the axle (between the bearings) doesn't deflect.

Dimensional Data	
Wheelbase	1100mm
Track Width	880mm
Ground Clearance	1.25inch
Length/Width/Height	1925.4mm/985.4mm/845.1mm
Kerb Weight	95 kg
Fuel Tank Capacity	5 litres

Table 1: Kart specification [1]

Cooling Type	Air cooled
No of Stroke	4 Stroke
No of cylinder	1
Displacement	124.7 cc
Maximum power	11.6 bhp @ 8000rpm
Maximum torque	11Nm @ 6500 rpm
Bore	52.4mm
Stroke	57.8mm
Compression ratio	9.2 : 1

Table 2: Engine Specifications [2]

Primary reduction	3.35 : 1
1st Gear	3.076 : 1
2nd Gear	1.944 : 1
3rd Gear	1.473 : 1
4th Gear	1.19 : 1
5th Gear	1.038:1
Final Reduction Ratio	3.071:1

Table 3: Gear Ratios [2]

Tires	Dimensions (in inches)
Front	10x4.5-5
Rear	11x7.1-5

Table 4: wheel size [3]

$$Number\ of\ teeth\ on\ driver\ sprocket = 14$$
$$Number\ of\ teeth\ on\ driven\ sprocket \cong 14 \times 3.071$$
$$Number\ of\ teeth\ on\ driven\ sprocket \cong 42$$

II. BENDING LOAD OF OUTER END OF REAR AXLE DUE TO ROAD GRADIENT

Assuming the road gradient of 4m raise in a height of 100m in length.

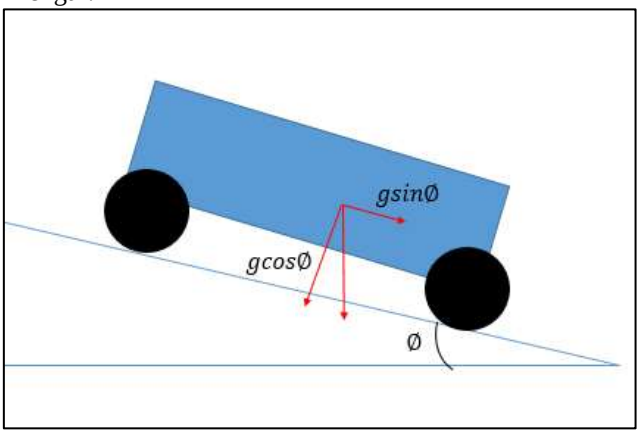


Fig. 1: Road gradient

$$\emptyset = tan^{-1}\frac{4}{100} \cong 2.3°$$

$$R1 = R + \frac{m \times H \times g \times sin\emptyset}{L}$$

All rights reserved by www.ijsrd.com

$$R1 = 509.6 + \frac{160 \times 0.2 \times 9.8 \times sin\emptyset}{1.1} = 521.04N$$

III. Bending Load on Outer End of Rear Axle Due to the Kart Acceleration

Here the kart acceleration is maximum at first gear. So the weight shift on rear axle is also maximum at first gear.
By considering the efficiency of the transmission system is 100%

$$T = Primary\ Reduction \times 1st\ Gear\ ration \times Final\ Reduction \times Engine\ Torque$$
$$T = 3.35 \times 3.076 \times 3.071 \times 11\ Nm$$
$$T = 348.1Nm$$

Driving force developed at the circumference of tire due to engine torque is,

$$Fr = \frac{T}{r} = \frac{348.1}{0.1397} = 2491.76N$$

By Assuming the coefficient of friction between tires and ground is 0.85, then available traction force during climbing is $2 \times R2 \times \mu = 2 \times 521.04 \times 0.85 = 885.76N$
Maximum possible acceleration, A.

$$A = \frac{2 \times R1 \times \mu}{m} = \frac{885.76}{160} = 5.53m/s^2$$

When kart is accelerating with maximum acceleration, A, then due to its inertial force that increases the loading on rear axle due to weight on rear side.

$$2 \times R2 = 2 \times R + \frac{mAH}{L}$$
$$R2 = 509.6 + \frac{160 \times 5.53 \times 0.2}{2 \times 1.1} = 590.03N$$

IV. Bending Load on Outer End of Rear Axle Due to Turning

When kart is turning then, it experiences a centrifugal force having the direction radially outside from the center of turn circle. And centrifugal force is equal to the centripetal force. Here the centripetal force is exerted due to the friction force between tires and ground surface. Hence the maximum possible value of centrifugal force is totally depends on the maximum available friction force between tires and ground. Weight shift on outer wheel is maximum, if centripetal force is equals to maximum static friction then.

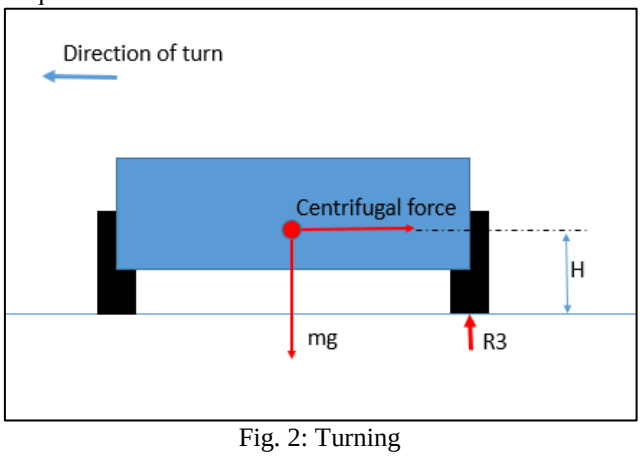


Fig. 2: Turning

$$Centripetal\ force = Maximum\ static\ friction$$

Assuming the coefficient of friction between tires and ground is 0.85, then, $Centripetal\ forcce = \mu \times m \times 9.8 = 0.85 \times 160 \times 9.8$

$$Centripetal\ forcce = Centrifugal\ force = 1332.8N$$

Taking moment about point of contact of inner wheel and ground

$P \times B = mg \times \frac{B}{2} + Centripetal\ force \times H$(Here P is the total loading on outer front and rear wheels during turning)

$$P = \frac{mg}{2} + \frac{Centripetal\ force \times H}{B}$$
$$P = \frac{160 \times 9.8}{2} + \frac{1332.8 \times 0.2}{0.88} = 1086.9N$$
$$R3 = \frac{P \times j}{L} = \frac{1086.9 \times 0.715}{1.1} = 706.5N$$

Here j is the distance of C.G of the kart from the from wheel axis
By superimposing the reaction forces R1, R2 and R3, we can get the combined resultant of maximum reaction force on rear axle. It is equivalent to that case, in which the vehicle is moving with acceleration A, climbing on road gradient of 4m raise in 100m length and simultaneously taking a turn. For design the rear axle to withstand in possible extreme scenario against bending.

$$R4 = R1 + R2 + R3$$
$$R4 = 521.04 + 590.03 + 706.5 = 1817.57N$$

V. Bending Load on Rear Axle Due to Traction Force to Propel the Vehicle

Driven force is developed at the circumference of the tire due to engine torque, in case of pure rolling, there is always equal and opposite friction force is acting at the point of contact between tire and ground to propel the vehicle. The effect of this driving friction force, causes the bending of rear axle shaft in a plane of parallel to the ground.

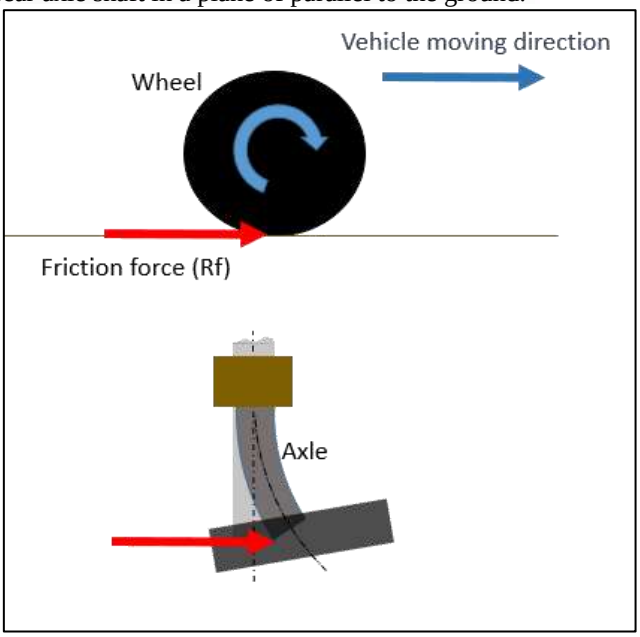


Fig. 3: Propulsion

While considering the maximum reaction force, R4 getting higher value of the available traction as compare to the maximum circumferential force, Fr develop due to engine torque.

All rights reserved by *www.ijsrd.com*

$$Rf = \mu \times R4 = 0.85 \times 1817.57 = 1545N$$

Thus Fr load is to be considered for maximum bending load parallel to the ground on the axle shaft.

$$Fr = 1248.1N.$$

VI. Resultant of Reaction Forces

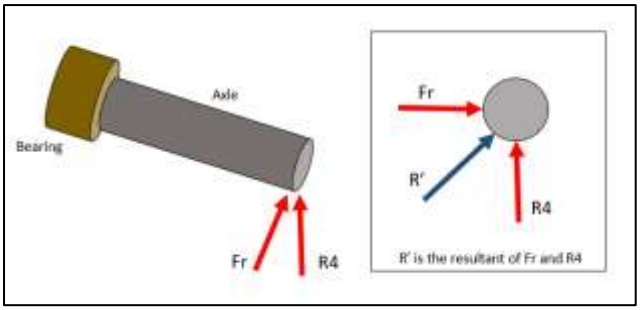


Fig. 4: Resultant Reaction R'

$$R' = \sqrt{Fr^2 + R4^2} = \sqrt{1248.1^2 + 1817.57^2} = 2204.83N$$

VII. Bending Stress Analysis ($\sigma 1$) due to the Resultant Load (R') at Wheel Hub Center

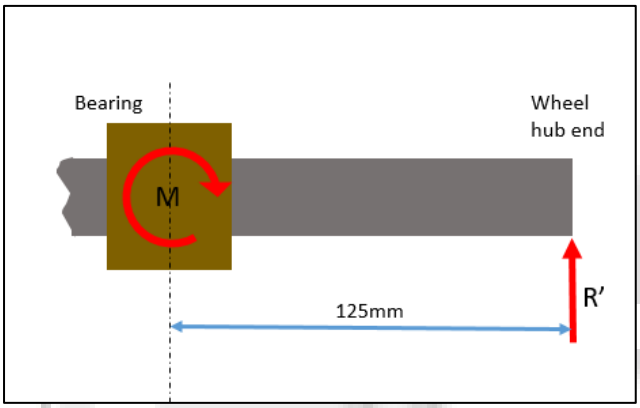


Fig. 5: Bending moment M1

$$M1 = R' \times 0.125 = 2204.83 \times 0.125 = 275.6Nm$$

$$\sigma 1 = \frac{M1}{Z} = \frac{275.6}{\frac{\pi}{32} \times D^3} N/m^2$$

VIII. Maximum Twisting Torque During Accelerating and Braking

In this case, force at the circumference of tire due to the engine torque is less than the available traction force on rear wheel during the maximum possible weight shift. Thus the possible max twisting moment that axle may have subjected is equal to engine torque. $Te = T = 348.1Nm.$

$$\tau e = \frac{Te}{Zp} = \frac{Te}{\frac{\pi}{16} \times D^3} = \frac{348.1}{\frac{\pi}{16} \times D^3} N/m^2$$

Maximum possible twisting moment during braking by considering 100% effective braking is due to maximum available traction, Rf on rear wheel. During braking some weight of the kart is shift towards the front end. Hence during this the reaction force at the rear end is always smaller than the reaction force R, at stationary condition. For calculating the twisting moment Tb during braking reaction force R is considered for conservative approach.

$$Tb = R \cdot \mu \cdot r = 509.6 \times 0.85 \times 0.3095 = 134.06Nm$$

$$\tau b = \frac{Tb}{Zp} = \frac{134.06}{\frac{\pi}{16} \times D^3} N/m^2$$

IX. Bending Analysis of the Inside Portion of the Axle Shaft Between Bearings Due to Engine Torque

For go-kart chain and sprocket mechanism is very commonly used to transmit the power from engine to the rear axle. Bending load due to engine torque on axle can be calculated by finding the resultant of all the forces exerted by chain on each engaged teeth of the sprocket of axle. And the magnitude of this resultant force is maximum when the angle of contact between chain and sprocket is 180+2α°, here α is the pressure angle of chain and sprocket.

Assuming 17° pressure angle for calculation of resultant force P1 on axle.

Here 25 teeth are in contact with the chain, then the tangential force at each teeth of the sprocket during transmitting the maximum torque Te at first gear is

$$\frac{Te}{25 \times radius\ of\ sprocket} = \frac{348.1}{25 \times 0.0645} = 215.87N.$$

And the resultant of these forces is calculated graphically by using AutoCAD.

Resultant bending load P1 is 2760N.

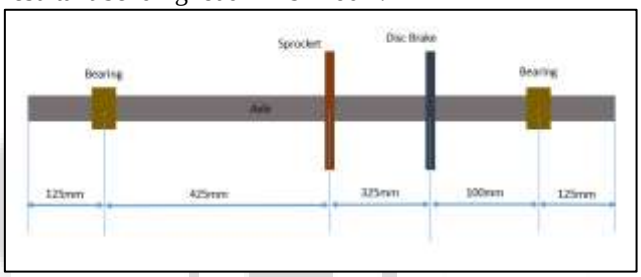


Fig. 6: Position of Sprocket, disc and Bearing center

Position of the sprocket is at the center of the rear axle.

$$M2 = \frac{2760}{2} \times 0.55 = 759Nm$$

$$\sigma 2 = \frac{M2}{Z} = \frac{759}{\frac{\pi}{32} \times D^3} N/m^2$$

X. Bending Load on Rear Axle Shaft During Breaking

The more appropriate size of disc brake is of 200mm in diameter because it provides enough clear distance of the brake disc from the ground. So that the rear disc of Apache RTR 180 is selecting for this go kart that is 200mm in diameter. And the effective radius of the disc is 90mm. For designing the rear axle during effective barking, by considering that the brake caliper exert the force on disc that develop resisting torque equal to Tb.

$$P2 = \frac{Tb}{0.09} = \frac{134.06}{0.09} = 1489.55N$$

I have fixed the position of disc at the distance of 750mm from left rear wheel bearing.

$$Reaction\ force\ at\ the\ right\ bearing, (Rb') = \frac{P2 \times 0.75}{0.85}$$

$$= \frac{1489.55 \times 0.75}{0.85} = 1314.30N$$

$$M3 = Rb' \times (0.85 - 0.75) = 1314.30 \times 0.1 = 131.43Nm$$

$$\sigma 3 = \frac{M3}{Z} = \frac{M3}{\frac{\pi}{32} \times D^3} = \frac{131.43}{\frac{\pi}{32} \times D^3} N/m^2$$

All rights reserved by *www.ijsrd.com*

XI. MATERIAL SELECTION

AISI 1020 is selected for rear axle of kart due to its availability and cheap as compare to AISI 4130. Lower yield strength and machinability is compromised with their cost.

Material	Tensile strength	Yield strength	Machinability
AISI 1020	420 MPa	350 MPa	65%

Table 5: Material Properties [6]

XII. COMBINED STRESS ANALYSIS ON THE REAR AXLE FOR VARIOUS FOLLOWING CASES

A. Case 1:- Analysis of the Outer Portion of Axle, i.e from Bearing to Hub Center.

When kart is accelerating with acceleration A, then the rear axle shaft may subjected to twisting moment due to engine torque and simultaneously to bending load due to weight shift during accelerating, climbing and turning. Considering a small element at the surface of the shaft where tensile stress due to bending and shear stress due to twisting moment are maximum.

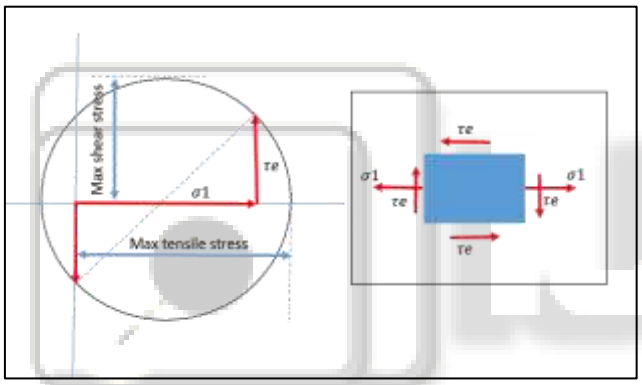


Fig. 7: Mohr's Circle

$$\sigma max = \frac{\sigma 1}{2} + \sqrt{\left(\frac{\sigma 1}{2}\right)^2 + \tau e^2} = \frac{3664.85}{D^3} N/m^2$$

$$\sigma min = \frac{\sigma 1}{2} - \sqrt{\left(\frac{\sigma 1}{2}\right)^2 + \tau e^2} = -\frac{857.61}{D^3} N/m^2$$

By using the shear strain energy theory

Taking factor of safety is 1.5 over yield strength 350(MPa).

Getting diameter of shaft is 26.12(mm)

B. Case 2:- Analysis of the Inner Part of Axle, I.E between the Bearings during Torque Transmission.

This possible cases of loading, during maximum torque transmission through chain and sprocket.

Considering a small element at the surface of the shaft where bending stress due to the resultant of forces act on each teeth of the sprocket by chain and shear stress due to twisting moment, Te are maximum.

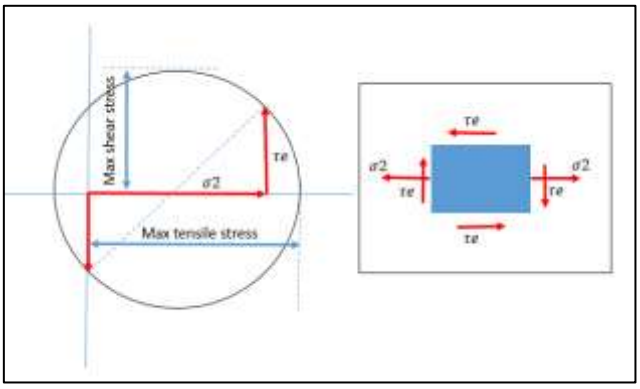


Fig. 8: Mohr's Circle

$$\sigma max = \frac{\sigma 2}{2} + \sqrt{\left(\frac{\sigma 2}{2}\right)^2 + \tau e^2} = \frac{8118.27}{D^3} N/m^2$$

$$\sigma min = \frac{\sigma 2}{2} - \sqrt{\left(\frac{\sigma 2}{2}\right)^2 + \tau e^2} = -\frac{387.15}{D^3} N/m^2$$

By using the shear strain energy theory

Taking factor of safety is 1.5 over yield strength 350(MPa).

Getting diameter of shaft is 32.91(mm)

C. Case 3:- Analysis of the Inner Part of Axle, I.E between the Bearings during Braking.

When brakes will apply, the shaft will experience a bending load, P1 and a twisting moment, Tb. Considering a small element at surface of the shaft where both bending stress and shear stress are maximum.

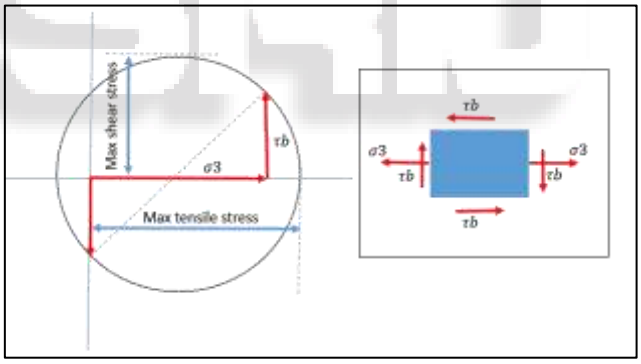


Fig. 9: Mohr's Circle

$$\sigma max = \frac{\sigma 3}{2} + \sqrt{\left(\frac{\sigma 3}{2}\right)^2 + \tau b^2} = \frac{1625.50}{D^3} N/m^2$$

$$\sigma min = \frac{\sigma 3}{2} - \sqrt{\left(\frac{\sigma 3}{2}\right)^2 + \tau b^2} = -\frac{286.78}{D^3} N/m^2$$

By using the shear strain energy theory

Taking factor of safety is 1.5 over yield strength 350(MPa).

Getting diameter of shaft is 19.70(mm)

D. Case 4:- Analysis of the Rear Axle for the Direct Shear Stress Caused by Various Lateral Loading on the Axle.

$$Direct\ shear\ stress = \frac{max(R', P1, P2)}{\frac{\pi}{4} \times D^2}$$

Shear yield strength = 0.577*tensile yield strength

All rights reserved by *www.ijsrd.com*

But the case of direct shear is very less significant as compare to other cases.

XIII. CONCLUSION

I studied various possible load conditions, as the rear axle might be subjected during operation. And calculated the appropriate diameter of the axle shaft by superimposing the possible load conditions in different scenario and selecting the maximum value of diameter 32.91(mm) say 33(mm) of the axle for safe function of the shaft in all cases during operation.

ACKNOWLEDGMENT

I am sincerely thankful to team Inferno DTU. Team Inferno DTU is a project based technical team of Delhi Technological University, founded in September 2013, with the sole focus on innovation. During working in team as a design person, I have learnt lot of things.

I am also Thankful to all my seniors of Inferno DTU who mentor me and giving me a right directions to go ahead.

I hope this research work helps to all those student who are going to participate in Go kart Championship.

REFERENCES

[1] Indian Karting Championship (IKC) Rulebook, Season 4, 2018
[2] https://bikeadvice.in/honda-cbf-stunner-125-review-sankesh/
[3] International Go kart Championship Rule Book 2017-18
[4] http://chain-guide.com/basics/2-1-2-engagement-with-sprockets.html
[5] https://bikeadvice.in/tvs-apache-rtr-180-ownership-review-pramod/
[6] https://www.azom.com/article.aspx?ArticleID=9145, for material properties
[7] PK Jain, Design, Development and Fabrication of a Go-Kart, Volume 6 Issue 2 (2018), 15 July 2018

ABBREVIATION

T	Maximum torque transmitted through axle, Nm
r	Radius of tyre, (m)
Fr	Force developed at circumference of wheel due to engine torque, (N)
μ	Coefficient of friction
m	Mass of kart with operator, (kg)
g	Gravity (9.8m/s^2)
H	Height of the C.G of vehicle, (0.2m)
j	Horizontal distance between front axle and C.G of vehicle. (0.715m)
A	Max acceleration at 1st Gear
L	Wheel base, (m)
B	Track width, (m)
R	Normal reaction force at each wheel of rear at static condition, (N)
R1	Normal reaction force on rear wheel due to road gradient, (N)
R2	Normal reaction force on rear wheel due to kart acceleration A, (N)
R3	Normal reaction force on rear wheel during turning of kart, N
R4	Resultant of all reaction forces (R1, R2, R3), N
Rf	Traction force
R'	Resultant of all reaction forces R1, R2, R3, Rf,, (N)
Te	Maximum twisting moment due to engine torque, (Nm)
Tb	Maximum twisting moment due to available traction force during braking, (Nm)
P1	Bending load on axle due to engine torque, (N)
P2	Bending load on axle during applying brake, (N)
M1	Bending moment due to load R', (Nm)
M2	Bending moment due to engine torque, (Nm)
M3	Bending moment due to applying brakes, (Nm)
τe	Maximum shear stress due to twisting moment Te, (N/m^2)
τb	Maximum shear stress due to twisting moment Tb,(N/m^2)
E	Young's modulus of Elasticity of material
D	Outer diameter of axle, (m)
I	Second moment of area,(m^4)
Ip	Polar moment of inertia, (m^4)
Z	Section modulus, (m^3)
Zp	Polar modulus, (m^3)
$\sigma 1$	Maximum bending tensile stress at extended portion of the axle
$\sigma 2$	Maximum bending tensile stress due to engine torque
$\sigma 3$	Maximum bending tensile stress due to applying brakes

All rights reserved by *www.ijsrd.com*

View publication stats

J. Inst. Eng. India Ser. B
https://doi.org/10.1007/s40031-020-00509-x

ORIGINAL CONTRIBUTION

Design and Control of Medium-Voltage Multilevel Converter for Direct Grid Integration of Photovoltaic System

Jyoti Kulkarni[1] · **Narendra Kumar**[1] · **Bhim Singh**[2]

Received: 13 February 2020 / Accepted: 21 October 2020
© The Institution of Engineers (India) 2020

Abstract This paper deals with the design and control of a medium voltage cascaded H-bridge (CHB) multilevel converter for the direct grid integration of a large photovoltaic (PV) system. It presents the design considerations for selecting an optimum number of levels in the multilevel converter, which are to be both efficient and economical. This system is designed for medium voltage, 11 kV, and hence it can be directly interfaced with the grid without a step-up transformer. It contributes large cost savings in the installation, operation and maintenance of the large solar plant. With the increased number of levels, the converter output voltage waveform is a good approximation to the sine waveform. Therefore, the total harmonic distortions (THDs) of the output AC converter voltage and the grid current are reduced, which significantly improves the power quality of the output voltage waveform reducing the size and hence the cost of the filter. The effect of an increased number of levels on THDs is also studied. The control employed is a decoupled control with the synchronous reference frame—phase-locked loop (SRF-PLL), and a 19-level cascaded converter is operating at the low switching frequency. The separate improved perturb and observation (P&O) maximum power point tracking (MPPT) algorithm is used for tracking the maximum power from each array of the multistring solar photovoltaic (SPV) system. Both steady-state and dynamic performances of the system are simulated in MATLAB/ Simulink environment and validated in accordance with the IEEE-519 standard. A substantial reduction in THDs of the output AC converter voltage and the grid current, with increased number of levels, is observed in this system.

Keywords CHB · THD · Solar PV Generation · MPPT · Improved P&O · Phase-Shifted PWM · Power Quality · PLL and SRF

Introduction

Today, the world is at crossroads for energy scenario. Every country wants to be developed. Per capita energy consumption is one key index to assess the progress of a nation. To be at par with other countries, developing nations are investing in power generation. Most of the investment is preferred in conventional energy sources. At the same time, the environmental challenges are mounting worldwide. The key challenge now is to balance progress and the environment. This is measured in terms of carbon footprint.

After the Paris agreement (International Solar Alliance) 2016, in the United Nations Framework Convention on Climate Change (UNFCCC), each signatory nation is bound to reduce its carbon footprint. The best way is to utilize the solar energy—the cleanest of all renewable energy sources (RESs). Out of this concern, an International Solar Alliance (ISA) has been established in 2016. The focus of ISA is to provide affordable solar power for every domestic household by the year 2030. India, for example, has set an ambitious target to have 100 GW installed solar capacity by the year 2022 and reduction in emission intensity by 33–35% by the year 2030 [1, 2]. In view of the above, one of the key challenges for the

✉ Jyoti Kulkarni
jyotikulkarni0803@yahoo.com

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

[2] Department of Electrical Engineering, Indian Institute of Technology, Delhi, India

Published online: 10 November 2020

engineers is the grid integration of RESs. Creating local pool and localization of solar generation at the point of use are other challenges [3].

The large-scale grid integration of solar power plants is to be the key focus area as the targets cannot be met without it. To be able to generate bulk solar power using the available solar panels, it requires series-parallel connection of solar panels to achieve desired current and voltage ratings [3]. Multilevel converters are offering such a facility to achieve desired power levels and grid integration. Out of the many multilevel topologies reported in the literature [4–6], for large PV applications, CHB topology is an exciting topology as it naturally needs a separate DC source for each H-bridge in CHB multilevel converter. Despite the increased hardware cost and complexity, multistring topology allows individual MPPT control of each string. It also provides increased modularity and improves power output from the solar module despite possible module mismatch and partial shading [7]. In the case of multilevel converters, the challenge is to generate a near sinusoidal voltage devoid of any harmonics. This is possible with an increase in switching levels. An increased switching level in CHB topology increases the complexity of the circuit and control techniques and also increased cost. Whereas on the other hand, it improves the power quality with a near sinusoidal waveform and also can reduce stress on individual switches, i.e., de-rating of the components, which can also reduce the cost of components [8–11]. A proper trade-off is required.

To extract maximum power from a PV panel, it is required to operate at a voltage and a current on the PV curve at the maximum power point. This can be achieved using the MPPT algorithm. With an appropriate MPPT algorithm and a switching strategy such as SRF-PLL, CHB multilevel converters are able to provide a good solution to the problem of grid integration of solar energy [12–14]. SRF control is used to control the grid current as it provides zero steady-state error with a good dynamic performance [15].

The phase-shifted PWM is the most suitable strategy for the cascaded inverter as the output voltage of individual H-bridges is almost identical, the rotation of switching patterns is not required and also switching devices operate at the same switching frequency and conduction time [16, 17].

Several single- and double-stage grid interfaced PV systems are reported in the literature [18, 19]. In a given system, a single-stage grid interfaced PV system is used as it reduces the cost of a DC-DC converter and also reduces the control complexity.

The multilevel converter provides the requisite medium and high voltage, so there is no need to provide an interfacing transformer for the grid. Only a low-frequency isolation transformer may be used for the grid integration in the case of single-stage PV system, or the galvanic isolation can be provided for the implementation of the control algorithm [20].

The improved P&O MPPT algorithm considerably reduces the ripple in the PV voltage; hence, it is used for tracking maximum power from the panel in the system discussed [21–23].

The main problem with a large SPV connected grid system is to design the voltage source converter (VSC) using high rating components with lower switching losses. Multilevel and multipulse converters are generally used to feed the large PV power into the grid [24–26]. Multipulse converters are used in a large solar PV grid-connected system to minimize harmonics in the voltage waveform resulting in good power quality, lower switching losses and reliability in comparison to a two-level topology but the good designing of a transformer and interfacing magnetic are required for proper operation of the multipulse-based system.

The system discussed for a medium voltage grid integration of PV system with the consideration of all design factors is a 19-level CHB multilevel converter grid integrated 1 MW power, 11 kV voltage, 50 Hz power frequency system with 500 Hz switching frequency and phase-shifted PWM. The number of levels in the multilevel converter in this system is selected with all design considerations like the cost of IGBTs, arithmetic and logical operations (ALOs), THDs and device voltage utilization factor (DVUF). The performance analysis of the system is carried out in MATLAB/Simulink.

System configuration

In high voltages (3.3, 4.5, 6.5 kV), costly IGBTs are available in the market but the lower voltage (0.6, 0.9, 1.2, 1.7 and 2.5 kV) IGBTs are comparatively lower in cost and have better technology. Hence, a cascaded connection of low voltage IGBTs for achieving medium voltage can be used for developing a low-cost inverter, and the designed system can facilitate a direct connection to the medium voltage AC network, with an improved output power quality. At the same time, this leads to a linear increase in the component number and control complexity. Hence, to realize a better performance versus price ratio, finding the optimum value for the number of levels is important. Usually, the 7-level to 21-level multilevel inverters for 11 kV are considered suitable for feeding the power at the given system voltage [8].

As the cost of a semiconductor device is significant for a medium voltage system, a higher DVUF is necessary for making the system design a cost-effective one. The DVUF

for an 11 kV system with a different number of levels is summarized in Table 1 [8].

In semiconductor devices, the inverter section losses are the sum of conduction losses and switching losses. As one increases the number of levels in a multilevel converter, the carrier frequency reduces and the active switching device count increases linearly. Hence, switching losses are reduced and at the same time conduction losses are increased. Moreover, the device voltage ratings determine an IGBT's on-state voltage drops and diode's forward voltage. For these reasons, total losses in the inverter section are almost constant and, therefore, despite the change in the number of levels, multilevel inverter's efficiency remains almost constant. Therefore, efficiency is not a consideration in the selection of levels while designing multilevel inverter.

Table 2 [8] gives the count of ALOs required in a switching section, the count of IGBTs for different levels, THD(%), and the cost in US$ for an 11 kV system for comparing the performance, cost and complexity of the system.

For comparing the variety of the inverter system on the basis of the cost, performance and control complexity comprehensively, a normalized index is defined as follows [8]:

$$I_{\mathrm{dw}} = (w - w_{\min})/(\, w_{\max} - w_{\min}) \tag{1}$$

where w is a scored value and $w_{\min}$ and $w_{\max}$ are the lowest and highest values of the indicator, respectively. The normalized index values for 11 kV are given in Table 3 [8].

For an 11 kV inverter, the value of the total index is the lowest for 19-level, so a 19-level is an appropriate choice for an 11 kV PV system. Beyond 19-level, power quality improvements and the reduction in semiconductor cost are negligible and it only increases parts (components) count and control complications as their increase is linear with the increase in a number of levels [9–11].

Table 1 DVUF in an 11 kV system for different levels [8]

Number of levels	Rated device voltage in kV	DVUF (%)
11	3.3	90
15	2.5	96
19	1.7	99
21	1.7	90

Table 2 An 11 kV system [8]

Number of levels	11	15	19	21
IGBTs	60	84	108	120
ALOs	55	77	99	110
THD(%)	8.55	6.00	4.27	4.25
Cost (US$)	142,080	198,912	8748	9720

Table 3 Normalized index for 11 kV system [8]

Number of levels	11	15	19	21
Performance	1.00	0.41	0.01	0.00
IGBT Cost (US$)	0.70	1.00	0.00	0.005
Complexity	0	0.17	0.5	1.00
Total Index	1.7	1.58	0.51	1.005

System design

Figure 1 shows a configuration of the system with a 19-level, 3-phase grid interfaced multilevel converter of power rating of 1 MW and 11 kV with 50 Hz power frequency and a switching frequency of 500 Hz. This PV system is connected to the grid through an interfacing inductor for reducing the ripple content in the current. Because of the low switching frequency, the requirement of the filter to filter the high switching frequency noise in the voltage is mitigated. As the PV system is designed for medium and higher power ratings, the SPV system is feeding power to the grid only for power transmission purposes.

In a cascaded multilevel converter, the number of levels selected is equal to (2 * number of H-bridge's per phase +1) [17]. So a 19-level CHB converter requires 9 H-bridges per phase and each H-bridge is fed by the separate PV array, which is easily available in a large SPV system. An improved P&O-based MPPT algorithm is used to extract the maximum power of the PV array. The variables measured for implementing the control algorithm are the grid voltages (V_{sabc}), the grid currents (I_{sabc}), the converter line voltages (V_{cabc}), the DC link voltage (V_{dc}), PV array voltage (V_{pv}) and PV current (I_{pv}). The designed system includes the design of the PV array, the DC link capacitor and the DC link voltage. Other design specifications of the system are given in Table 4.

Selection of V_{dc} voltage

In the case of m number of level converters, for each individual H-bridge connected to the PV array, the DC-link voltage (V_{dc}) is calculated as [17],

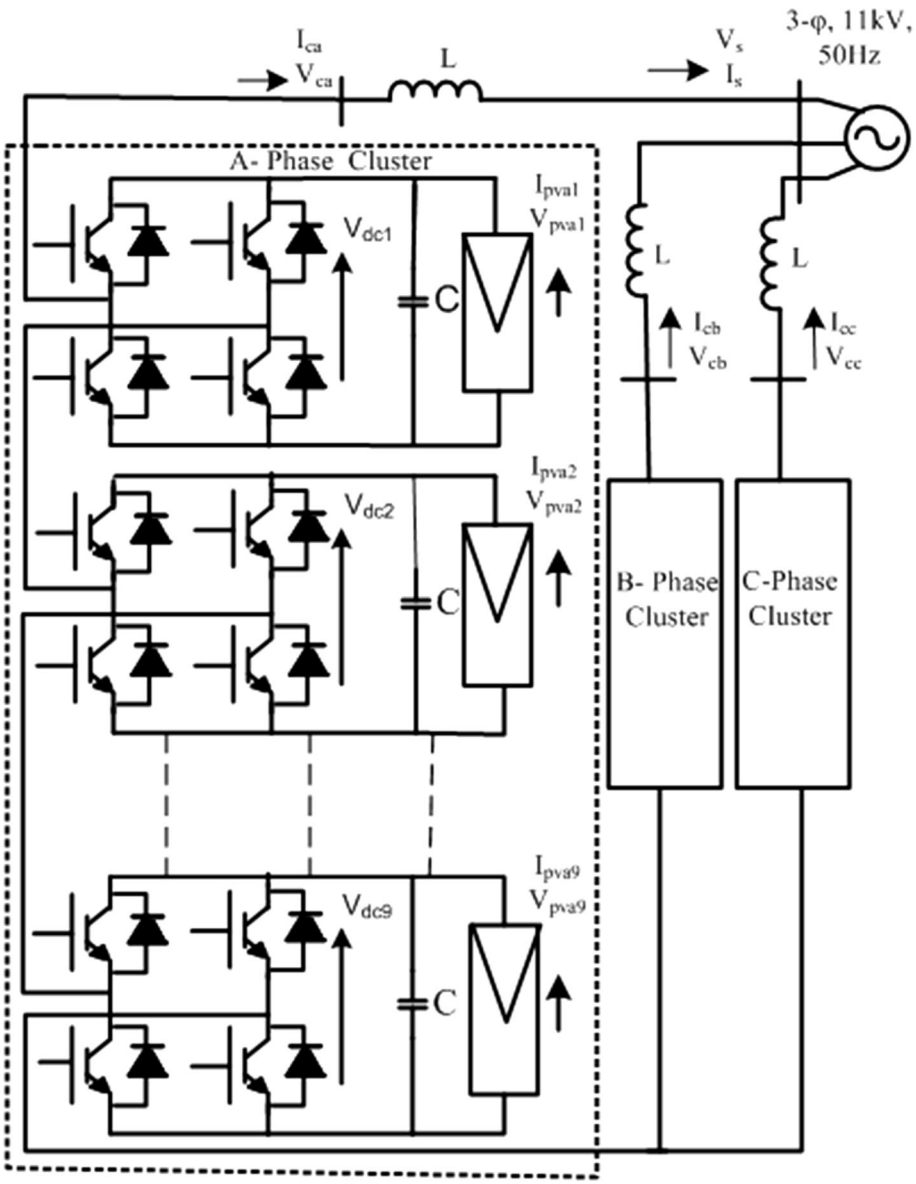


Fig. 1 A 19-level grid interfaced converter PV system

$$V_{\text{rms}} = 0.612 * (m-1) * V_{\text{dc}}$$
$$V_{\text{dc}} = 998.5V \quad (2)$$
$$V_{\text{dc}}\text{is taken as } 1010\ V.$$

Design calculations and selection of PV array

The power rating of each PV array for a 19-level CHB converter consisting of total 27 PV arrays (9 per phase) for a 1 MW power grid interfaced system is as follows:

$$P_{\text{mmp}} = \text{Total power } (P)/\text{No. of PV arrays} = 1 * 10^6/27 = 37.04kW. \quad (3)$$

To get the required specifications P_{mmp}, I_{mmp} and V_{mmp}, Sun Power SPR-245NE-WHT-D model is selected from the PV array library, and the number of PV modules parallel and series are calculated as [25, 26].

$$N_{\text{s}} = V_{\text{dc}}/V_{\text{mmp}} = 1010/40.5 = 25 \quad (4)$$

$$N_p = P_{\text{mmp}}/\left(N_s * I_{\text{mmp}} * V_{\text{mmp}}\right) = 1 * 10^3/(25 * 6.05 * 40.5) = 6 \quad (5)$$

Design of capacitor for DC-link

The DC-link capacitance (C_{dc}) for each H-bridge is calculated as per a principle of conservation of energy for the DC-link voltage (V_{dc}), with a recovery in 5 ms, a safety factor (k_1) 0.1, and a 1.2 overloading factor (a), as [27],

Table 4 Component specifications for a 19-level system

Component	Value
Power	1 MW
Voltage (V_{rms})	11 kV
Frequency	50 Hz
Switching Frequency	500 Hz
Inductor (L)	4 mH
No. of parallel paths in SPV array (N_p)	6
No. of series paths in SPV array (N_s)	25
Voltage of PV array at maximum power (V_{mmp})	40.5 V
Current of PV array at maximum power (I_{mmp})	6.05 A
Open circuit voltage of PV array (V_{oc})	48.8 V
Short circuit current of PV array (I_{sc})	6.43 A
DC link capacitance (C_{dc})	2000 μF
DC link voltage (V_{dc})	1010 V

$$\frac{1}{2} * C_{dc}\left(V^2{}_{dc}\text{--}V^2{}_{dc1}\right) = K_1 * 3 * V * a * I * t$$
$$C_{dc} = \left(0.1 * 37.04 * 10^3 * 1.2 * .005*\right)/ \left(0.5 * (1010 * 1010\text{--}998.5. * 998.5)\right) = 1924.35\mu F \quad (6)$$

It is selected as $C_{dc}(C)$ = 2000 μF.

Control algorithm

The algorithm for the control of the SPV system integrated to the three-phase ac grid is shown in Fig. 2.

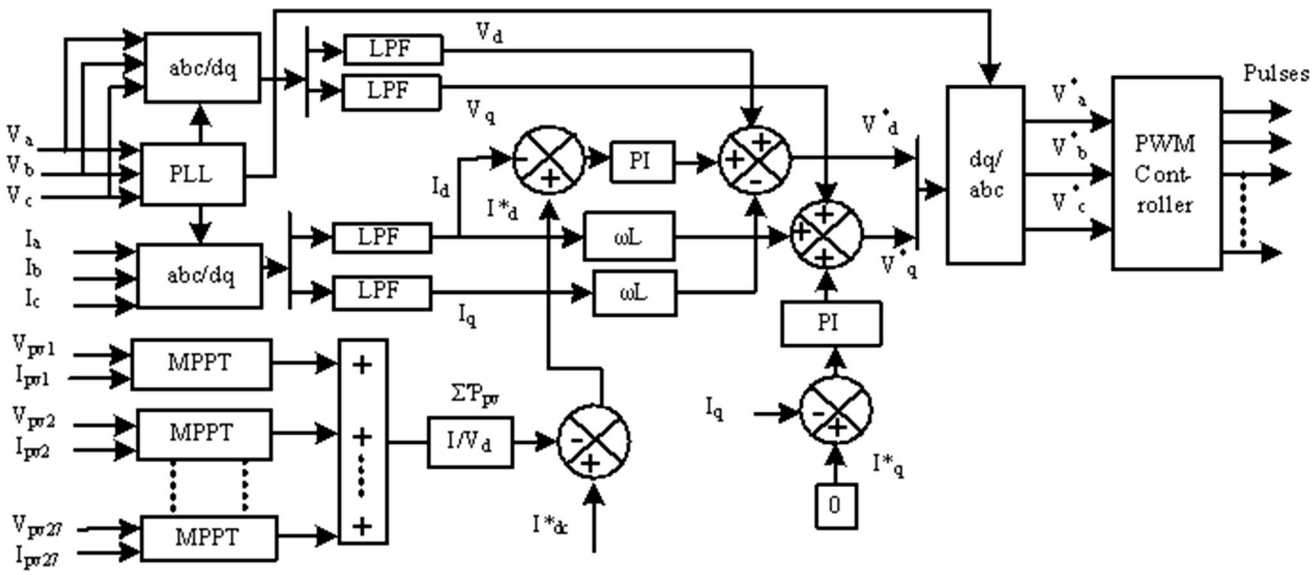


Fig. 2 Control algorithm

The maximum power from the array which is tracked using converter voltages and currents is necessitated. A decoupled current control having a feed-forward term, the control for the DC-link capacitor, synchronous reference frame (SRF) and phase-locked loop (PLL) for synchronization are used for controlling VSC [13, 14].

MPPT algorithm

The maximum power is extracted from the PV array, and a suitable V_{dc_ref} is generated by using the MPPT algorithm. Out of many reported MPPT algorithms, the improved P&O algorithm is preferred owing to its superiority vis-à-vis other algorithms. In the literature [22, 23], it is reported that the P&O algorithm is the PV array independent, it is a true MPPT, an implementable in both digital and analog form, devoid of periodic tuning requirement and easy to implement. As far as an improved P&O algorithm is concerned, apart from the benefits of the P&O algorithm, it also takes care of the drift effect owing to increased insolation, by taking into account the change in the current apart from the change in voltage and power. The MPPT governing equations are as follows:

$$V_{pv} * (k) = V_{pv} * (k-1) + \Delta V_{pv}, \text{if } dP_{pv} < 0 \text{ and } dV_{pv} < 0 \text{ or } dP_{pv} > 0, dV_{pv} > 0 \text{ and } dI_{pv} < 0 \quad (7)$$

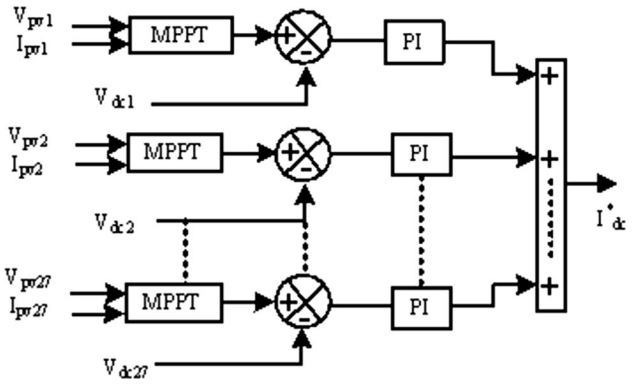


Fig. 3 Control algorithm for generating I^*_{dc}

$$V_{pv}*(k) = V_{pv}*(k-1) - \Delta V_{pv}, \text{if } dP_{pv} < 0 \text{ and } dV_{pv} > 0 \text{ or } dP_{pv} > 0, dV_{pv} > 0 \text{ and } dI_{pv} > 0 \text{ or } dP_{pv} > 0 \text{ and } dV_{pv} < 0 \quad (8)$$

Control scheme for VSC

The voltages obtained from the measurement of each DC link (V_{dc}) and given by each MPPT algorithm (V_{dc_ref}) are compared, and the error voltage (V_e) between the two is processed by a proportional-integral (PI) controller as shown in Fig. 3, which gives the improved P&O MPPT algorithm. To integrate the solar PV power generating system with the grid synchronization of the grid voltages and currents with the value of I^*_d, the reference DC-link current, expressed at k and (k-1) instant as follows

$$I^*_d(k) = I^*_d(k-1) + K_p\{V_e(k) - V_e(k-1)\} + K_i * V_e(k) \quad (9)$$

where K_p and K_i are the proportional and integral gain constants for the PI controller.

The sum of all reference DC-link currents along with the feedforward term is compared with the direct current component obtained by the Park transformation from the grid currents [27],

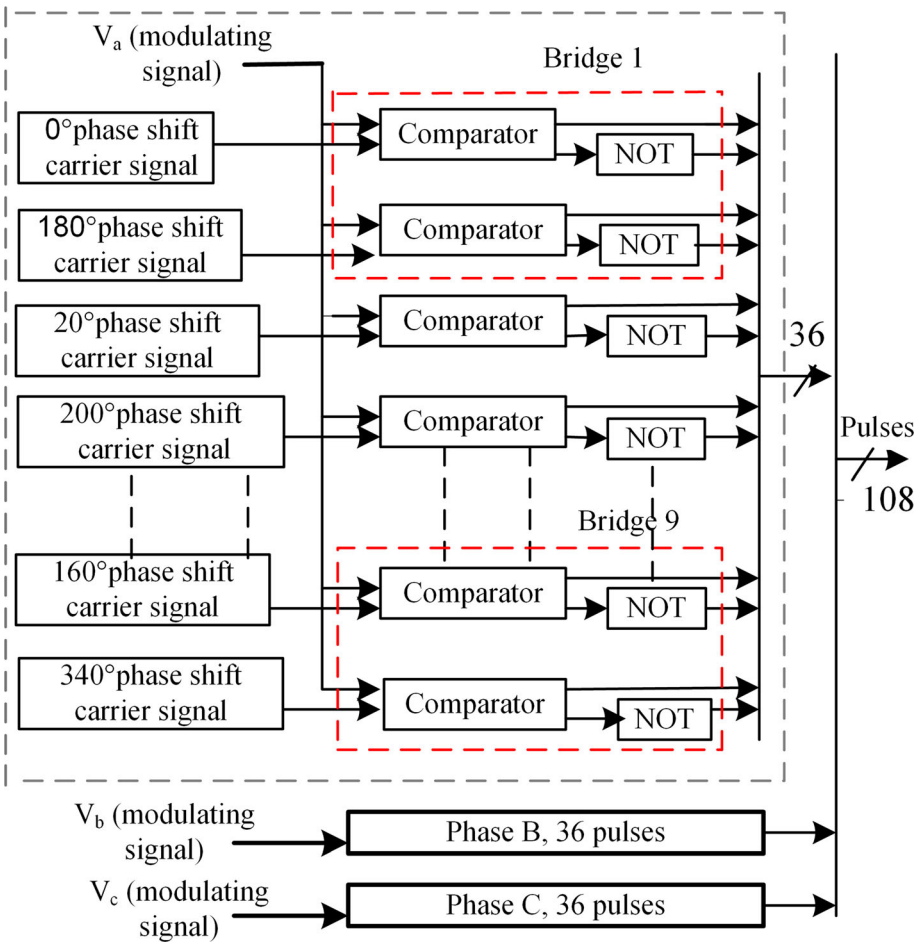


Fig. 4 Scheme for generation of PWM signal

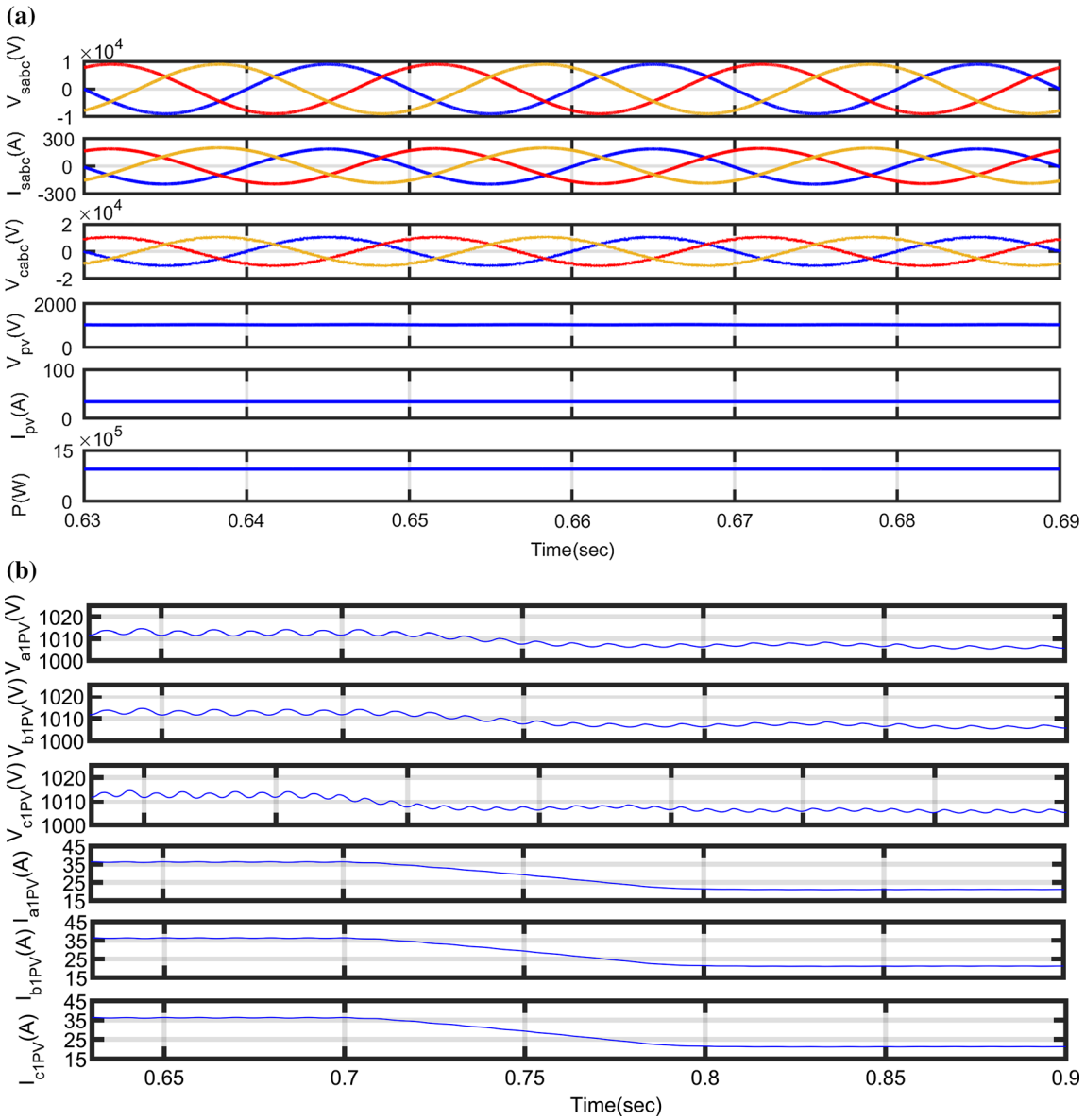


Fig. 5 (**a**) Steady-state performance at insolation level 1000 W/m^2 and temperature 25 °C (**b**) Dynamic performance with insolation level changed from 1000 to 500 W/m^2 at 0.7 s

$$\begin{bmatrix} I_d \\ I_q \\ I_0 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \sin(\theta) & \sin(\theta - 2\pi/3) & \sin(\theta + 2\pi/3) \\ \cos(\theta) & \cos(\theta - 2\pi/3) & \cos(\theta + 2\pi/3) \\ 1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} V_d \\ V_q \\ V_0 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} \sin(\theta) & \sin(\theta - 2\pi/3) & \sin(\theta + 2\pi/3) \\ \cos(\theta) & \cos(\theta - 2\pi/3) & \cos(\theta + 2\pi/3) \\ 1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (11)$$

The grid detection angle (θ) for the Park transformation is obtained from a three-phase PLL. I^*_q the reference quadrature current (set to zero for feeding the grid current at unity power factor) is compared with the quadrature current component obtained from the Park transformation.

Similarly, the direct (V_d) and the quadrature (V_q) components of the grid voltages are calculated using the Park transformation as follows.

The reference direct-axis and quadrature-axis voltages at instants k and (k-1) are estimated as follows:

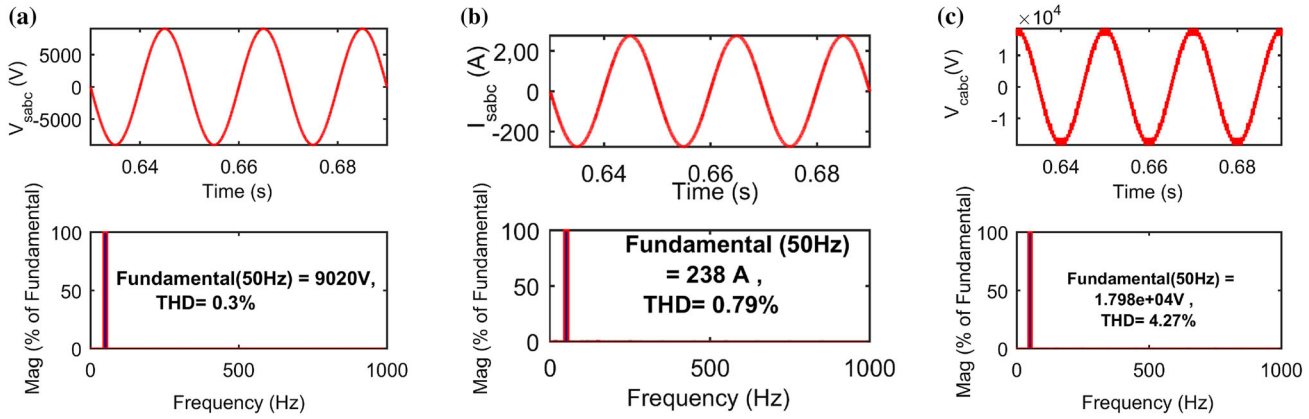


Fig. 6 Waveform and harmonic spectra of 19-level and 11 kV system (**a**) line voltage (**b**) line current (**c**) converter line voltage at 1000 W/m^2 and 25 °C

$$V^*_d(k) = V^*_d(k-1) + K_{p2}\{I_{de}(k) - I_{de}(k-1)\} + K_{i2} * I_{de}(k) + V_d - I*_q\omega L \tag{12}$$

$$V^*_q(k) = V^*_q(k-1) + K_{p3}\{I_{qe}(k) - I_{qe}(k-1)\} + K_{i3} * I_{qe}(k) + V_q + I*_d\omega L \tag{13}$$

where $I_{de} = (I^*_d - I_d)$, $I_{qe} = (I^*_q - I_q)$ and ωL is the coupling reactance. K_{p2}, K_{i2}, K_{p3} and K_{i3} are the respective proportional and integral gains of the direct axis PI current controller and the quadrature axis PI current controller.

PWM control

A multicarrier PWM technique is most widely used in multilevel converter. The phase-shifted PWM technique with a low switching frequency (500 Hz) is used for modulation. For a 19-level CHB, a total of 27 H-bridges are used; hence, a total of 108 pulses are generated as shown in Fig. 4. The phase shift between the carrier signals in the case of CHB is calculated as 360°/ (L-1), where 'L' denotes the number of levels. For a 19-level, the phase shift between the carrier signals is 20° [17].

Results and discussion

A 19-level CHB for an 11 kV grid interfaced PV system with an optimum selected number of levels is modeled and analyzed in the MATLAB environment, and the results obtained are shown in Figs. 5-6. The values of converter line voltage THD obtained for a 19-level converter are also compared with those of 7-level and 11-level systems.

Response of system

The steady-state response of the system is studied at the insolation level of 1000 W/m^2 and temperature 25 °C, and the results are depicted in Fig. 5(a). Simulated results consisting of the grid voltages (V_{sabc}), the grid currents (I_{sabc}), the AC line voltages of VSC (V_{cabc}), SPV voltage (V_{pv}), SPV current (I_{pv}) and power (P) are obtained to validate the system performance. As the grid voltage and the grid current are in phase in steady-state response, the maximum active power transfer is evident at a power factor of unity. The dynamic response as shown in Fig. 5(b) is studied with a constant temperature of 25 °C and an insolation change from 1000 W/m^2 to 500 W/m^2 at 0.7 s. Simulated results consist of the voltage and the current for a PV array 1 of A-phase (V_{a1pv} & I_{a1pv}), B-phase.

(V_{b1pv} & I_{b1pv}), C– phase (V_{c1pv} & I_{c1pv}). In dynamic response, with the decrease of the insolation level, the remarkable change in the PV current is observed, while there is no significant change seen in the PV voltage.

Improvement in power quality

Figure 6 shows the waveforms for the line voltage (V_{sabc}), the line current (I_{sabc}) and the converter line voltage (V_{cabc}) and harmonic spectra along with THD levels at 1000 W/m^2 and 25 °C temperature levels. The THD values of 0.30%, 0.79% and 4.27%, respectively, are observed for the line voltage, the line current and the converter line voltage for a 19-level grid interfaced 11 kV SPV system. The supply current THD is well within limits set by the IEEE-519 standard with reduced size of the filter [28]. The waveform and harmonic spectra for 7 and 11-level multilevel converters are also shown in Fig. 7.

Table 5 summarizes the results of the converter line voltages THDs for 7, 11, and 19 levels at 1000 W/m^2 insolation and 25 °C temperature levels, given in Figs. 6

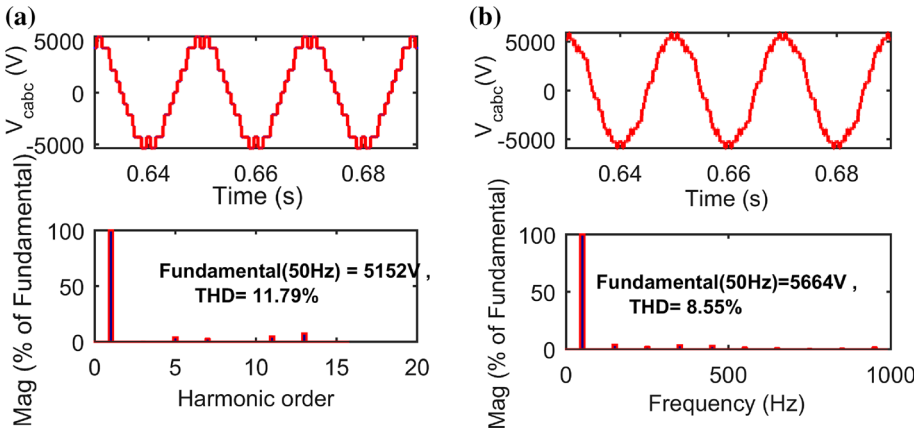


Fig. 7 Waveform and harmonic spectra (**a**) 7-level with 3.3 kV system (**b**) 11-level with 3.3 kV at 1000 W/m^2 and 25 °C

Table 5 Converter voltage THDs comparison for various levels

Level	THDs (%)
7	11.79
11	8.55
19	4.27

and 7. The results prove that the THD is reduced with increased number of levels.

Conclusion

The modeling and design of a medium voltage grid interfaced cascaded multilevel converter for PV system with a selected number of levels have been carried out with decoupled current controller based on SRF-PLL and phase-shifted PWM at low switching frequency for medium and large-scale PV direct grid integrated plant. The reduction in the size of the line filter due to reduced THDs and an elimination of step-up transformer from the traditional system for integration with the grid contribute to large cost savings in installation, operation and maintenance of such large solar plants. The cost of IGBTs and voltage stress on each switching device are reduced with an optimum number of level. The low-switching frequency employed reduces the switching losses and lowers the acoustic noise, which has also contributed in the reduction of filter size and cost. The converter line voltage THDs for different number of levels is compared. The improved power quality of the system owing to reduced THDs is observed. The system performance is found satisfactory as per the IEEE standard 519–1992 [28].

References

1. Earnest and Young. The Paris agreement; What it means for India" [online]. Available from https://www.ey.com/Publication/vwLUAssets/EY-the-paris-agreement-2017/$FILE/EY-the-paris-agreement.pdf
2. 107 Sunshine countries… [online]. Available from https://web.archive.org/web/20151208071012/http://liveindia.in/107-sunshine-countries-may-shape-pms-vision
3. F. Blaabjerg, R. Teodorescu, M. Liserre, A.V. Timbus, Overview of control and grid synchronization for distributed power generation systems. IEEE Trans. Industrial Electronics **53**(5), 1398–1409 (October 2006)
4. J. Rodriguez, J.S. Lai, F.Z. Peng, Multilevel inverters: a survey of topologies, controls and applications. IEEE Trans. Ind. Elect. **49**(4), 724–738 (2002)
5. S. Kouro, M. Malinowski, K. Gopakumar, J. Pou, L.G. Franquelo, B. Wu, J. Rodriguez, M.A. Perez, J.I. Leon, Recent advances and industrial applications of multilevel converters. IEEE Trans. Ind. Elect. **57**(8), 2553–2580 (2010)
6. H. Akagi, Classification, terminology and application of the modular multilevel cascaded converter (MMCC). IEEE Trans. Power Electron. **26**(11), 3119–3130 (Nov 2011)
7. A. Noman, K. Addoweesh, A. Alabduljabbar, A. Alolah, Cascaded H-Bridge MLI and Three-Phase Cascaded VSI Topologies for Grid-Connected PV Systems with Distributed MPPT. Int. J Photoenergy **2019**, 7642919 (2019)
8. R. Islam, Y. Gua, J. Zhu, A multilevel medium-voltage inverter for step-up-transformer-less grid connection of photovoltaic power plants. IEEE J Photovoltaics **4**(3), 881–889 (May 2014)
9. M. Rabiul Islam, Y. Guo, J.G. Zhu, D. Dorrell, Design and comparison of 11 kV multilevel voltage source converters for local grid based renewable energy systems. *IECON 2011 – 37th Annual Conf. of IEEE Industrial Electronics Society, Melbourne, VIC* pp. 3596–3601, (2011)
10. M.R. Islam, Y.G. Guo, J.G. Zhu, H-bridge multilevel voltage source converter for direct grid connection of renewable energy systems. *IEEE PES Innovative Smart Grid Technologies*, Perth, WA, 1–7 (2011)
11. T. Kerekes, E. Koutroulis, D. Sera, R. Teodorescu, M. Katsanevakis, An optimization method for designing large PV plants. IEEE J Photovoltaics **3**(2), 814–822 (April 2013)
12. S. Kouro, B. Wu, A. Moya, E. Villanueva P. Correa, J. Rodriguez, Control of a cascaded H-bridge multilevel converter for grid connection of photovoltaic systems. *35th Annual Conf. of IEEE Industrial Electronics, Porto* pp. 3976–3982 (2009)

13. B. Singh, M. Kandpal, I. Hussain, Grid integration of single stage SPV system using asymmetric cascaded 7-level VSC, *39th National Systems Conf. (NSC)*, Noida, pp. 1–6 (2015)
14. M. Mehtani, I. Hussain, M. Khandpal, B. Singh, Integration of multiple source SPV grid tied system using 27-level cascaded H-bridge converter, *2016 IEEE 7th Power India Inter.Conf. (PIICON)*, Bikaner, (2016)
15. Y. Han, H. Chen, Z. Li, P. Yang, L. Xu, J.M. Guerrero, Stability Analysis for the Grid-Connected Single-Phase Asymmetrical Cascaded Multilevel Inverter With SRF-PI Current Control Under Weak Grid Conditions. IEEE Trans. Power Electronics **34**(3), 2052–2069 (March 2019)
16. B.P. McGrath, D.G. Holmes, Multicarrier PWM strategies for multilevel inverters. IEEE Trans. Ind. Elect. **49**(4), 858–867 (August 2002)
17. Bin Wu, High-Power Converters and AC Drives, IEEE, 2006
18. C. Jain, S. Goel, B. Singh, A distribution grid tied multifunctional SPV system operating with control approach based on decoupled adaptive neural network, *IEEE Industry Applications Society Annual Meeting*, Addison, TX, pp. 1–8 (2015)
19. A.K. Verma, B. Singh, D. T. Shahani, Grid interfaced solar photovoltaic power generating system with power quality improvement at AC mains, *IEEE Third Inter.Conf. on Sustainable Energy Technologies (ICSET),* Kathmandu, pp. 177–182 (2012)
20. R. Gonzalez, E. Gubia, J. Lopez, L. Marroyo, Transformerless single-phase multilevel-based photovoltaic inverter. IEEE Trans. Ind. Elect. **55**(7), 2694–2702 (July 2008)
21. B. Subudhi, R. Pradhan, A comparative study on maximum power point tracking techniques for photovoltaic power systems. IEEE Sus. Energy **4**(1), 89–98 (Jan 2013)
22. J. Ahmed, Z. Salam, A modified P&O maximum power point tracking method with reduced steady state oscillation and Improved tracking efficiency. IEEE Transactions on Sustainable Energy **22**(2), 1506–1514 (Oct 2016)
23. A.K. Singh, I. Hussain, B. Singh, An improved P&O MPPT algorithm for single stage three-phase grid integrated solar PV system, *IEEE 7th Power India Inter. Conf. (PIICON), Bikaner,* pp. 1–6, (2016)
24. I. Hussain, B. Singh, Investigations on solar PV grid interfaced power generating system using two-level twelve-pulse double photovoltaic inverter. IEEE Trans. Ind. Elect. **55**(7), 2694–2702 (July 2008)
25. I. Hussain, B. Singh, Investigations on solar PV grid interfaced power generating system using two-level twelve-pulse double bridge converter, *9th Inter. Conf. on Industrial and Information Systems (ICIIS),* Gwalior pp. 1–6 (2014)
26. B. Singh, I. Hussain, Grid integration of single stage solar PV power generating system using 12-pulse VSC, *IEEE 6th India Inter. Conf. on Power Electronics (IICPE),* Kurukshetra, pp. 1–6 (2014)
27. B. Singh, A. Chandra, K. Al-Hadded, *Power Quality: Problems and Mitigation Techniques* (John Wiley & Sons, Inc., 2015)
28. IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems, *IEEE Standard* 519–1992 (9 April 1993)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Dynamic Voltage Frequency Scaling in Multi-core Systems using Adaptive Regression Model

Manjari Gupta
Department of Electronics and Communication Engineering
Malaviya National Institute of Technology
Jaipur, India
2015rec9032@mnit.ac.in

Lava Bhargava
Department of Electronics and Communication Engineering
Malaviya National Institute of Technology
Jaipur, India
lavab@mnit.ac.in

Indu Sreedevi
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
s.indu@dce.ac.in

***Abstract*—A learning-based manager that controls the power budget through dynamic voltage frequency scaling (DVFS) in a multi-core processor has been proposed in this paper. The core statistics are collected and employed to predict the next interval power consumption and are thereby used to determine the best suited voltage-frequency setting for each core. The aim is to maximize performance while containing the power consumption per-core. The presented solution is realized in Snipersim and the fine-grained DVFS algorithm is included through Python scripting. Simulation results demonstrate that the proposed approach is able to achieve 6.6 % energy-reduction and average power-savings of 27.4% against the existing state-of-the-art algorithm (Steepest Drop) for various allocation schemes.**

***Index Terms*—Dynamic voltage frequency scaling, energy and performance tradeoff, Multicore processors, Regression model**

I. Introduction

A single CPU with multiple execution cores which have the ability to work in parallel on distinct operations is termed as a *multicore system*. Power administration and performance augmentation in the MPSoC is a fundamental customer expectation. Energy-efficient systems are highly desirable in various fields like mobile platforms as it reduces the number of battery recharges considerably. This also improves the chip's reliability and lifetime. The overall power consumed by a processor can be split into dynamic and leakage power. Charging and discharging of capacitative inputs generates dynamic power while static power consumption due to leakage current is the reason for Leakage power. Frequency scaling provides quadratic energy savings as: E α Capacitance X voltage^2 X frequency X cycles [1]. Though activity reduction and lowering the clock speed curtails the power consumption linearly, it has been observed that the change is infinitesimal for a fixed number of cycles. It has been observed that voltage scaling along with management of clock frequency gives a significant route for energy savings. [2] Modern microprocessors such as Intel XScale, Transmeta Crusoe, and AMD Athlon come with in-built DVFS capability.

The main goal of DVFS is dynamic scaling of the supply voltage of each core so as to facilitate "just enough" circuit speed to the core for processing system workload with sufficient speed while accommodating the throughput constraints and minimizing the energy dissipation simultaneously [3]. However, the effectiveness of this technique varies with granularity. Per-core DVFS gives high flexibility but requires a large number of voltage control knobs. In per-chip DVFS, all the cores are controlled by a single knob but diminish the extensibility. The proposed work will be working at core-level, hence providing fine-grained power manager.

The paper organization is as follows. Section 2, details recent work in DVFS along with machine learning (ML) algorithms. It also highlights the contributions made by this paper. Section 3, elaborates the methodology, framework and algorithm used. The experimental setup and result analysis are detailed in Section 4. Finally, section 5 concludes the work.

II. Related Works and Paper Contribution

Various techniques of determining power management policies have been studied in recent times. Developing an energy-efficient system has been a long standing requirement in our day-today activities such as survillence system [21] and IoT [22]. A cluster-based DVFS technique is proposed in [4] to reduce power consumption. The solution aims to find an intermediate solution to the shortcomings of both per-chip and per-core DVFS. Thus, cores operating on same voltage levels are grouped together and DVFS is applied per-cluster. This technique handles both power overhead and reduces temperature variation among cores. Genetic algorithm based DVFS is applied in [5] to determine static task mapping. The task assignment is done to minimize power along with improving core aging and transient fault tolerance. DVFS combined with other strategies like task migration [6] has also been explored to develop a system which is able to achieve multi-objective optimization criteria. The efficiency of any search-based strategy diminishes as the number of cores inflates. Winter et al. in [7] highlights these scalability limitations using a directed local search method - Steepest Drop. Though the method aims to solve a high-performance power efficiency problem, it fails to include power budget overshoot in the scope.

In the recent past, machine-learning techniques have been explored for energy and power consumption minimization. Techniques like DVFS and DPM have been thoroughly studied

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

and have proved to be effective in targeting power reduction. However, run-time analysis of most energy efficient voltage/frequency level is a challenging task and requires accurate predictions in a short amount of time. Thus, the capability of learning based algorithms when combined with the effectiveness of these methods has attracted academia. Dhiman et al. in [8] presented a dynamic power management technique (DPM) for a single-core processor which is based on an online learning model. A single DPM policy is dynamically selected from a set of existing values using an ML algorithm. The algorithm aims to converge to a predetermined frequency/voltage setting. C. Maury et al. developed a dynamic system, AC-TOR (Adaptive Concurrency Throttling Optimization Run-time System) in [9] to control energy-levels while maintaining performance using Artificial Neural Networks (ANN). The work highlights the convenience and improvements achieved by using ANN over pre-existing algorithms. The model was an extension to previous work where multiple linear regression was used. Unlike our work, the linear regression model calculates performance predictors and not estimated power statistics. A Bayesian – classifier motivated supervised learning based power management has been studied in [10]. The framework takes a per-core approach to determine power management action on the basis of workload and focuses on the reduction of overhead associated with continuous monitoring of manager. Ye et al. in [11] proposed a Q-learning algorithm based DPM technique. Since the method covers the task allocation in the learning phase, it utilizes the idle period of cores to reduce DVFS scaling time. Our paper makes the following contributions :

- The proposed research work devise an *offline* adaptive regression - based power manager which learns the best-suited voltage-frequency (v-f) setting from amongst a set of predefined values such that performance is maximized under energy constraints.
- The control manager is queried during *online* operation to predict the optimal v-f setting yielding highest performance within power constraint. The manager uses already stored performance counter (IPC) and power statistics (Core power) to make the decision.
- The solution is implemented in Sniper simulator [12] and the regression model in a python script. The proposed method has been compared with Steepest Drop for benchmarks from PARSEC-suite with varying workloads in a system with as many as 32 cores and improves energy efficiency by up to **6.6%** . It is also demonstrated the inclusion of an efficient heap-sort that further improves the execution time and power-savings.

III. Methodology

A. Overview

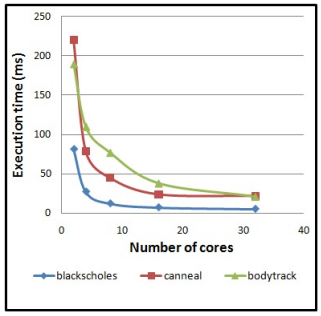
This work addresses *power-performance* tradeoff in multi-core systems. The effect of both power and performance optimization is still unexplored for learning based techniques. Additionally, the existing techniques suffer from scalability issue. In our work, computationally intelligent power manager using a hybrid approach has been used to *dynamically* distribute the power while maximizing the performance of the system. An efficient sort- system is used to handle the scalability issue. Instructions-per-cycle (IPC) represent instruction-level parallelism and memory access pattern of an application. IPC along with the base frequency dictate the performance capability of the application on the core. The operating frequency has a direct impact on the energy consumption of the core. Thus, altering the DVFS levels intelligently influences both IPC and Power component of the core.

B. Benchmark Selection

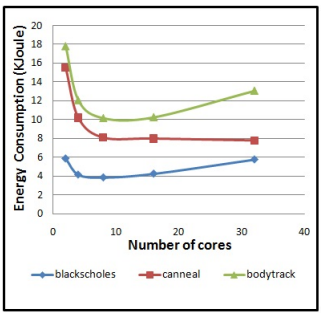
The proposed method uses benchmarks from PARSEC-suite which is industry standard applications [13]. The three benchmarks *(blackscholes,bodytrack, canneal)* are chosen such that they represent a varying working set and granularity. The proposed research work is focused on benchmarks that represent parallel, multi-threaded applications with low communication requirement. The details of the benchmark characteristics are provided in Table I [14]. The behavior of the benchmarks was studied for various frequency levels to understand their effect on the proposed power manager. Figure 1 (a) and 1 (b) provide a snapshot of the amount of variation in the energy spectrum and performance behavior across all the benchmarks. The execution time is higher for benchmarks with large amount of data usage as incur communication cost. While the energy-consumption largely depends on number of instructions (working set) & granularity.

TABLE I
CHARACTERISTICS OF PARSEC BENCHMARKS [13]

Benchmark name	**Working set**	**Granularity**	**Data Usage**		**Application domain**
			sharing	*exchange*	
blackscholes	small	coarse	low	low	Financial Analysis
bodytrack	medium	medium	high	medium	Computer Vision
canneal	large, unbounded	fine	high	high	Engineering



(a) Performance behavior



(b) Energy Consumption comparison

Fig. 1. Analysis of benchmarks at 1800 MHz

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

C. Data Collection

Core statistics for each benchmark executed on different combinations of v-f levels and number of active cores are collected through data-collection script in Python. This data symbolizes different core behavior and ensures that the predictive model is unbiased towards a particular pattern. As a part of data analysis, following core states for each core are extracted and stored in a csv file for further processing:

1. Core IPC (Instructions per cycle)
2. Core Power
3. Core Number
4. Operating frequency of each core

The power reported is a sum of dynamic & static power measured in Watts. The data is grouped as per operating frequency. Distinct predictive models are generated for each frequency in order to maximize efficiency. This core statistics forms the baseline for all implementations. For each benchmark, 62 data points are obtained per frequency. The total data set of 1,116 ($62 * 6 * 3$) points is divided into training & testing data in 7:3 ratio.The data matrix would look like :
$(Core\#, IPC, CorePower) : \{i, IPC_i, Pow(i, Vf_i)\}$
where $0 < i \leq 32$ and $Pow(i, Vf_i)$ is the power of core *i* at voltage / frequency level Vf_i.

D. Regression-based DVFS

The power manager should have knowledge of both hardware and software to control their behavior. A Linear regression model which uses system architecture and per-core statistics to control operational frequency has been presented in this paper. Regression models are being used to estimate continuous value of the dependent variables ($IPC_{predicted}$ and $Power_{predicted}$) for the next interval as a function of performance monitor inputs (number of cores, frequency,$IPC_{lastinterval}$ and $Power_{lastinterval}$).

Linear regression is simple supervised learning method which helps to identify a relationship between unknown variables(dependent) and known features (independent). The dependent variables are observed and are controlled through independent variables [15]. The relationship is defined through the below-mentioned equation:

$$h_\theta(x) = \alpha + \sum_{i=1}^{n} \theta_i x_i + \in \quad (1)$$

$h_\theta(x)$: $IPC_{predicted}$ or $Power_{predicted}$
α: intercept
$\in$: Error
x_i: vector of independent input values
θ_i: weights/coefficients calculated during the LR modelling
n: number of predictors
The efficiency of the predictive model is computed on the basis of Root Mean Square Error which is defined as :

$$J(\theta) = \frac{1}{m} \sqrt{\sum_{j=1}^{m} (h_\theta(x^{(j)}) - y^{(j)})^2} \quad (2)$$

m: number of training samples
$x^{(j)}$: jth value of predictor variables
$y^{(j)}$: jth measured value
$h_\theta(x^{(j)}$: jth predicted value
The order of complexity for the regression-based modelling for single-core is $O(mn^2+n^3)$ and multi-core system is $O(\frac{mn^2}{p} + \frac{n^3}{p} + n^2 log(p))$ [16] where p is number of cores. Thus, a fast model can be achieved for small number of features/variables (n) and learning samples (m).

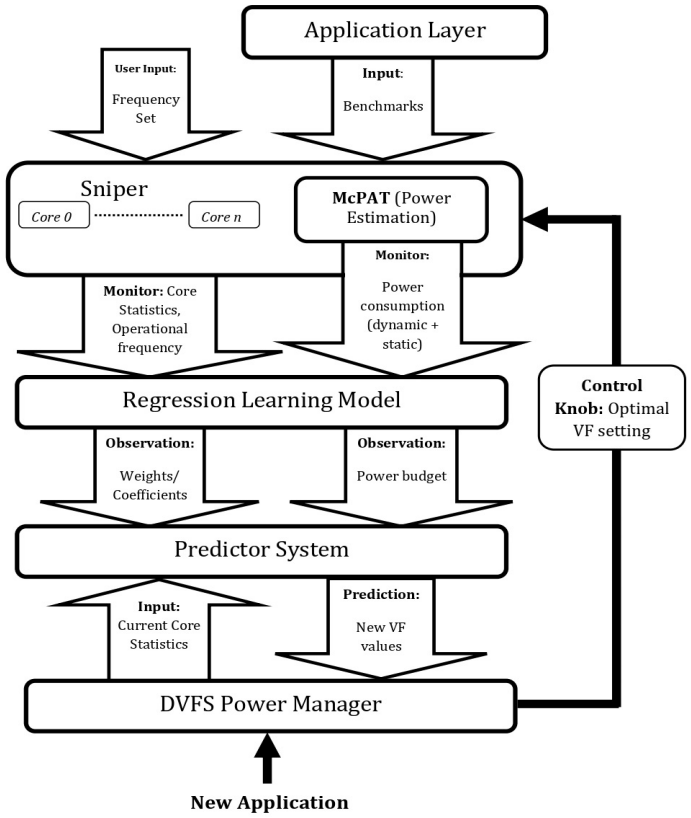


Fig. 2. Overview of Framework

E. DVFS Strategy

The overview of the framework is mentioned in Fig. 2 . The no-DVFS data collected is used as a guideline to compare the effectiveness of the proposed methodology. The statistics collected for the highest number of cores and maximum frequency combination is used for over-budgeting and the model is able to predict and prevent them. A per-core power manager is implemented through python script and will monitor the core behavior. Once the learning is done, this manager decides to increase or decrease the frequency of the core depending on power budget and other DVFS trigger criteria. Although our emphasis is on per-core dynamic DVFS, it also considers the global power management by considering the overall chip power budget.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

Once the performance and power models have been generated, the power budget is calculated for each frequency. The highest power consumed by a core for the lowest frequency is designated as per-core power budget. The global chip threshold is calculated by multipying it with number of active cores. The residual power budget *(delta budget)* is calculated as the difference between the power budget and the sum of current power dissipation by all the cores. This value signifies the available power for DVFS management. However, instead of assigning the surplus power on a random basis, where it make use of heap sort for a structured allotment. The *delta budget* for each core is arranged in both Max and Min heap-sort to ensure the core with maximum/minimum headroom for frequency change is identified. The manager then verifies that the current power utilization by the core should be less than per-core power budget. If this check fails, the manager sends lower frequency settings to the simulator. If the check passes, then the manager validates the following two conditions:
1) the predicted power for the next interval ($Power_{predicted}$) should be less than per-core power budget:

$$Power_{predicted} < \frac{Powerbudget}{Numberofcores} \tag{3}$$

2) current IPC should be greater than the predicted IPC i.e.

$$IPC_{current} < IPC_{predicted} \tag{4}$$

If both the conditions are met, the DVFS manager increases the frequency of the core. The optimal VF setting is then communicated to the simulator. The simulation statistics are updated after every delta sim time, set as 500 μs [17]. The power consumption for both Max-sort and Min-sort have been compared and discussed in section 4.

The advantage of using heap sort is two-fold. First, the method removes the randomness in the process of power allocation and helps to quickly identify the core with the maximum/minimum scope of frequency scaling. Second, in case the core fails any condition of frequency variability, the system does not need to recalculate the statistics for all the remaining cores and can simply pick up the next child-node from the heap. In worst case, $n * p$ iterations for n cores and p v-f levels are needed to reach the highest frequency level. However, an efficient heap data structure only requires $O(p.n.logn)$ time to attain the highest level, thus reducing the overall execution time.

Algorithm 1 Pseudo code for LR-DVFS method

1: **Input :** i , *IPC current (i)* , *Pow (i, VF_i), N,* $0 < i \leq 31$
2: **Output :** Opt_VF_i
3: **Variable :** *delta budget* $_i$
4: **calculate** IPC $_{predicted}$, Power $_{predicted}$ for each core
5: *power budget* ← *Pow (i, VF_i = lowest) * N*
6: *delta budget* $_i$ ← *power budget* - $\sum_{i=0}^{N-1} Pow\ (i, VFi)$
7: built a max – heap of delta budget based on each core
8: **while** *delta budget* $_i$ > 0 **do**
9: Pop the max-heap and get tuple (*i, Pow (i, VF_i))*
10: **if** *Pow (i, VF_i)* < $\frac{power\ budget}{N}$ **then**
11: **if** (Power $_{predicted}$ < $\frac{power\ budget}{N}$) and (*IPC current (i)* < IPC $_{predicted}$) **then**
12: Opt_VF_i ← Increase frequency level
13: **end if**
14: **else**
15: Opt_VF_i ← Decrease frequency level
16: **end if**
17: **recalculate** *delta budget* $_i$
18: **end while**

IV. Experimental Evaluation and Results

A. Experimental setup

The effectiveness of the suggested technique is substantiated using Sniper simulator with PARSEC benchmark suites. The simulated configuration is summarized in Table II. Sniper gives the requisite core statistics (IPC, Power, Operational Frequency) at run-time for varying workload benchmarks. Additionally, Sniper supports DVFS implementation with an acceptable scale of flexible v-f levels and provides power estimation through McPAT. Intel's McPAT integrates power and area model calculated over a period of time for MPSoCs and is used to obtain a precise power state of the chip [18]. The v-f settings mentioned have been extracted from the voltage/frequency specifications of the Xeon 5550 processor [19] and have been set as a table in dvfs.py script in Sniper. However, the lowest frequency chosen is 1000 MHz due to McPAT constraint. A data-collection program and the power manager are created in Python script. The algorithm followed is mentioned for reference. The steps 7 - 18 are repeated for min-sort.

In an ideal scenario, DVFS instantaneously transitions from one frequency to another. However, in reality, switching frequency takes time due to internal PLL (phase lock loop) locking time and capacitances [3]. Existing DVFS controllers consider the latency required for transition as negligible which, if properly incorporated can improve the quality of frequency change decisions. The Intel Xeon 5550 processor has 70 μs switching latency for 4 cores at 1600 MHz [20]. The CPU voltage/frequency scaling time has been comfortably accommodated in the DVFS decision epoch (500 μs) to minimize its effect. However, this epoch may be longer than necessary for the lesser number of cores.

TABLE II
Simulated Architecture Configuration

Number of cores	2, 4, 8, 16, 32
Architecture	Intel Processor Gainestown
Technology node	45nm
L1-I/D Cache	32KB, 4-way,LRU
L2 Cache	256 KB, 8-way, LRU
L3 Cache	8 MB, 16-way, LRU
Frequency levels (GHz)	1 , 1.4, 1.8, 2.1, 2.7, 3.3
Voltage (V)	0.8 , 0.9, 1.1, 1.2, 1.4, 1.6
DVFS epoch	500 μs

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

B. Results and Analysis

1) Model Analysis: The target Power and IPC for the benchmarks are established by executing the baseline system with no prediction and no DVFS. These values imitate normal communication between the system architecture and the operating system. The power manager continuously monitors the states of all the active cores and evaluates the combination of voltage/frequency levels which will give the maximum throughput while keeping the overall chip power consumption under the global power budget. In order to obtain a more realistic solution, the core power budget is not considered as a constant value and is calculated by varying the number of cores. Our estimation of per core power budget for a single core under different frequencies is shown in Table III. The tabulated values have been extrapolated for different core statistics. The actual power range for the Xeon 5550 is not available for more than 4 cores. The data obtained from our method can be compared with actual power range which is 95W for 4-cores [19]. This calculation has been used as a starting point. A snapshot of transient IPC and

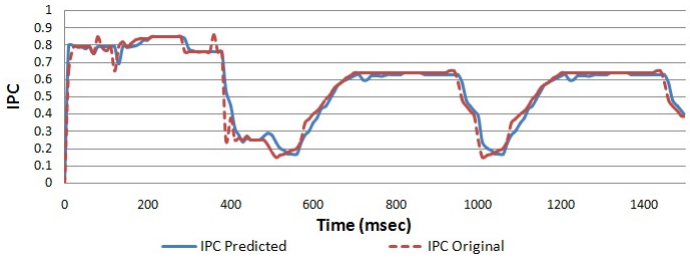


Fig. 3. Performance Model

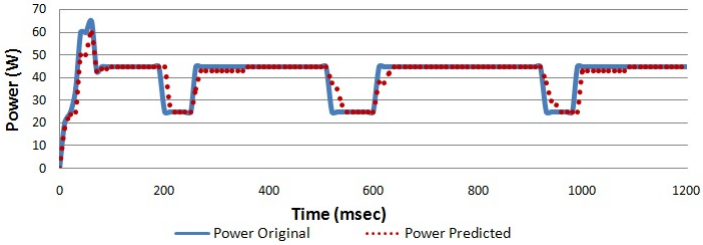


Fig. 4. Power Model

TABLE III
POWER THRESHOLD AND RMSE

Frequency level (MHz)	Power threshold (Watt)	Root Mean Square Error	
		IPC	Power
1000	2.3	0.71	0.455
1400	4.2	0.78	0.968
1800	11.6	0.89	2.589
2100	18.9	0.92	3.454
2700	23.7	0.96	4.865
3300	31.5	1.03	9.45

Power diagram for 2100 MHz for 4 cores is shown in Figure 3 and 4 respectively. The performance model compares the *predicted IPC* with *observed* IPC values. Similarly, the power model captures the accuracy of *observed* and *predicted* power values. It is observed that although both the models are able to fairly predict the new values, they are unable to capture transitions and sudden changes. The model is able to predict IPC meticulously as the IPC range is fairly small. Whereas, the power range for the different number of active cores working across varying frequencies is large. Also, the power consumption model shows large divergence as the number of cores increases. The accuracy of a linear model is judged by the Root Mean square error (RMSE) (2). For better understanding, the proposed work has tabulated RMSE for each model across different frequencies in Table III. It is evident that although the performance model predicts next interval IPC with good accuracy, the power model gives larger deviation for higher frequencies. The proposed research work aim to minimize this by using either a non-linear or a probability-based model in future work.It is also essential that the case of underprediction or overprediction are addressed in our method. If the current IPC is sufficiently less than the predicted value, then the frequency of the core is increased to meet the performance goal. Although, if the IPC or Power value is miscalculated, there will be a small amount of performance degradation. This issue can be handled in the next epoch by limiting the negative impact on performance. However, overprediction does affect the reduction in power consumption.

2) Power Consumption Analysis: Since our main aim is to achieve an energy-efficient system, the performance of our method is evaluated with Steepest Drop algorithm for maximum power savings. The base execution involves running the benchmarks without DVFS and regression model and the power consumed is noted as $Power_{normal}$. The power savings for our method is compared with the Steepest Drop for a system of 4,8,16,32 cores. Since Steepest Drop has not been implemented for 2 cores, hence the comparison has been excluded. For our method, ($Power_{modelscheme}$) corresponds to Power consumption for each of the below-mentioned schemes after linear modeling:

1) Sequential power budget allocation (no sort)
2) Max-heap sort (max-sort)
3) Min-heap sort (min-sort)

The metric computed to analyze Power savings is :

$$\frac{mean(Power_{normal}) - mean(Power_{modelscheme})}{mean(Power_{normal})} * 100 \quad (5)$$

Power savings for all the models and Steepest Drop is calculated according to equation (5) and plotted with respect to the number of cores. From Figure 5, it is evident that our method combined with max-heap sort constantly outperforms Steepest Drop.

Although the proposed method has an insignificant impact to 2-core structure, it gives promising results with a rise in the number of cores. For a quad-core system, Steepest Drop performs better than no-sort and min-sort scheme as they suffer from higher iterations in DVFS levels. The regression-based model surpasses Steepest Drop for 16, 32 core architecture.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

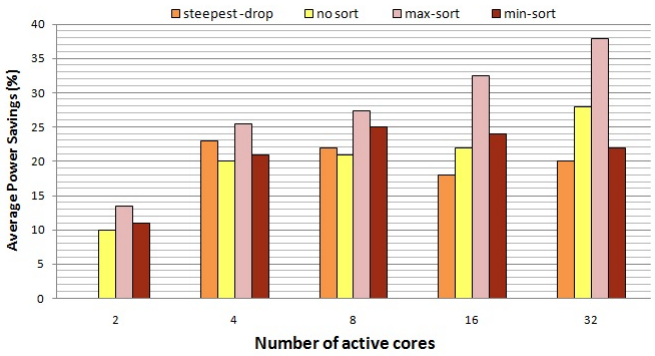


Fig. 5. Power-savings across different configurations

The two power allocation schemes - no sort and min - sort behave in a similar manner and show equivalent average power savings of 20.2% and 20.6 % respectively. However, min-sort suffers from more transitions as the same core is selected multiple times. This gap in the assignment scheme diminishes the efficiency for a higher number of cores. The max- sort, on the other hand, performs much better than both the remaining power assignment techniques and gives an average of 27.4% reduction in power consumption. Our LR-based power manager is also able to hold the overall chip power utilization under threshold levels. However, Steepest Drop exceeds the power budget more than 43% as it initially starts with the maximum frequency setting.

On the basis of the results, the following observations are obtained: 1) The focus of Steepest Drop is reducing the execution time rather than power dissipation. It achieves the same by curtailing the number of transitions in frequency. However, our method considers both power and IPC to achieve power-performance efficiency; 2) Steepest Drop initiates the system with highest power setting and then steps down the frequency. This condition makes the system less power-efficient; 3) Further, Steepest Drop takes the decision on the basis of the current state regardless of the next state. Since the states change frequently and Steepest Drop cannot predict the next state, the best action calculated on the basis of current state may result into higher power consumption.

V. Conclusion

Our work provides a meticulous control on the energy-performance agreement through the regression-based model which uses performance counters along with power estimates efficiently. The method aims to predict the next interval power and performance counters and use them to handle frequent changes in a structured manner. Our presented approach is able to achieve 20.2% , 20.6% and 27.4% average power savings as compared to a no-DVFS, no-regression technique and is able to improve average energy-efficiency by 6.6% for a 32-core system.

References

[1] S. Ghiasi, J. Casmira, D. Grunwald, “Using IPC variation in workloads with externally specified rates to reduce power consumption”, Complexity-Effective Design at ISCA27, June 2000.
[2] G. Dhiman, T. S. Rosing, “Dynamic voltage frequency scaling for multi-tasking systems using online learning”, Proceedings of International Symposium on Low Power Electronics and Design, pp. 207-212, 2007.
[3] K. Choi, R. Soma, M. Pedram, “Fine-Grained Dynamic Voltage and Frequency Scaling for Precise Energy and Performance Tradeoff Based on the Ratio of Off-Chip Access to On-Chip Computation Times”, IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, vol. 24, no. 1, pp. 18-28, Jan. 2005.
[4] T. Kolpe, A. Zhai, S. S. Sapatnekar, “Enabling improved power management in multicore processors through clustered DVFS”, Design Automation & Test in Europe Conference & Exhibition (DATE) 2011, pp. 1-6, 2011.
[5] A. Das, A. Kumar, B. Veeravalli, C. Bolchini, A. Miele, “Combined DVFS and mapping exploration for lifetime and soft-error susceptibility improvement in MPSoCs”, Proc. Int. Conf. Design Autom. and Test in Europe, pp. 1-6, 2014.
[6] K. Cao, J. Zhou, M. Yin, T. Wei, M. Chen. “Static Thermal-Aware Task Assignment and Scheduling for Makespan Minimization in Heterogeneous Real-Time MPSoCs”, International Symposium on System and Software Reliability (ISSSR), 2016.
[7] J. A. Winter, D. H. Albonesi, and C. A. Shoemaker, “Scalable thread scheduling and global power management for heterogeneous many-core architectures,” in PACT. ACM, 2010, pp. 29–40.
[8] G. Dhiman, T.S. Rosing. “Dynamic Power Management Using Machine Learning” , IEEE/ACM International Conference on Computer Aided Design, 2006.
[9] M. Curtis-Maury, K. Singh, S. McKee, F. Blagojevic, D. Nikolopoulos, B. De Supinski, M. Schulz, “Identifying energy-efficient concurrency levels using machine learning”, IEEE International Conference on Cluster Computing, pp. 488-495, 2007.
[10] E.-Y. Chung, L. Benini, G. D. Micheli, “Dynamic power management using adaptive learning trees”, Dig. Tech. Papers IEEE/ACM Int. Conf. Computer-Aided Design, pp. 274-279, 1999.
[11] R. Ye, Q. Xu, “Learning-Based Power Management for Multicore Processors via Idle Period Manipulation”, IEEE TCAD, 2014.
[12] T. E. Carlson, W. Heirman, and L. Eeckhout, “Sniper: Exploring the level of abstraction for scalable and accurate parallel multi-core simulations”, in International Conference for High Performance Computing, Networking, Storage and Analysis (SC), Nov. 2011.
[13] C. Bienia, S. Kumar, J. P. Singh, and K. Li, “The parsec benchmark suite: Characterization and architectural implications,” in Proceedings of the 17th international conference on Parallel architectures and compilation techniques, ACM, 2008, pp. 72 - 81.
[14] C. Bienia, S. Kumar, J. P. Singh, and K. Li, “The PARSEC benchmark suite: Characterization and architectural implications,” in Proc. 17th Int. Conf. Parallel Archit. Compilation Techn., 2008, pp. 72 - 81.
[15] J. Cohen, P. Cohen, S. West, and L. Aiken, “Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences”, Taylor & Francis, 2013.
[16] C.T. Chu, S.K. Kim, Y.A. Lin, Y. Yu, G.R. Bradski, A.Y. Ng, K. Olukotun, “Map-Reduce for Machine Learning on Multicore”, Proc. 20th Ann. Conf Neural Information Processing Systems (NIPS ’06), pp. 281-288, 2006.
[17] Z. Chen, D. Marculescu, “Distributed reinforcement learning for power limited many-core system performance optimization”, Proc. IEEE/ACM Design Autom. Test Eur. Conf. Exhibit. (DATE), pp. 1521-1526, Mar. 2015.
[18] S.Li,J.H. Ahn, R.D. Strong, J.B. Brockman, D.M. Tullsen,N.P. Jouppi.“McPAT: an integrated power, area, and timing modeling framework for multicore and manycore architectures”, 42nd Annual IEEE/ACM International Symposium on, pp. 469-480. IEEE, 2009.
[19] http://ark.intel.com/products/37106
[20] A. Mazouz, A. Laurent, B. Pradelle, W. Jalby, “Evaluation of CPU frequency transition latency,” Computer Science - Research and Development, 2014.
[21] Sungheetha A, Sharma R. Cost Effective Energy-Saving System in Parking Spots. Journal of Electronics. 2020;2(01):18-29.
[22] Nirmal D. Artificial Intelligence Based Distribution System Management and Control. Journal of Electronics. 2020;2(02):137-47.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:57:18 UTC from IEEE Xplore. Restrictions apply.

Volume 8 Issue 2 (2020) 99-102 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Effect of temperature on wear rate and coefficient of friction of Railway Brake Block

Vishal Kumar, Nishant Yadav, Prem Kumar, Roop Lal
Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India.

Article Info
Article history:
Received 15 February 2020
Received in revised form
29 May 2020
Accepted 8 June 2020
Available online 15 June 2020

Keywords: Wear Rate, Coefficient of friction, Railway Brake Block.

Abstract: Braking is a very complex process in overall railway assembly, braking system has a great importance and contribution towards the safety. During braking a number of phenomenon occurred like mechanical, thermal, electrical etc. which led to the high forces act on the brake blocks of railway results in thermal cracks, reduced life of blocks. It is estimated that a minimum of 60 brake shoes or brake blocks are used in the railway wagons which have to be changed within a month. In our study, author investigated and analyzed the braking tribology characteristics i.e. to determine the wear rate and coefficient of friction at a constant frequency at different temperature by keeping the stroke length constant. The experiment reveals the rate of wear and coefficient of friction varying at ambient and in summers during to continuous exposure of brake block for a fixed distance moved. The experiment is performed on the liner reciprocating tribometer by using stainless steel pin without any aid of lubricating material.

1. Introduction

A railway brake is a type of brake used to de-acceleration, control the acceleration at time of uphill or downhill or to stop for the wagons of the railway trains. The basic principle of railway braking is same as vehicles running on the road while the operational features are more complex because of multiple linked carriage system and also effectively used to stop at the same time.

In the earliest days of railway, braking technology was primitive. The first train had brakes operative on the locomotive tender and they were called as potters or brakeman to stop the train. Some railways fitted a special deep-noted brake whistle to locomotives to indicate to the porters the necessity to apply the brakes. At this stage, all the brakes were applied by operation using the screw and linkage to brakes. The braking effort achievable here was limited and also unreliable as it was fully dependent on the listening on the brakeman. Nowadays we use a frictional type of brake in which a fixed surface (shoe block) is brought into contact with moving parts to slow down or stop the vehicle named as railway brake shoes or railway brake pads. The process is known as shoe braking (or treads brake) .Brake shoe is a device for braking the moving group of cars and other type of rolling stock. In addition to the main task of braking the cars, the shoes are used to prevent the cars to move freely.

Shoe braking is obtained by direct pressure of brake blocks on the wheel tread. It is typical for freight trains and frequent for metro and suburban trains, either as unique braking system or coupled with other ones. Shoes braking power is limited and only applied when the speed is below 120km/hr.The application of shoe brake is raised in recent years due to its environment friendly coritents.

In shoe braking, brake block is pressed against the steel rim by the application of pneumatic pressure. During braking it is taken into consideration to employ continuous, constant (unfluctuating) and exact pressure so that brake remains into contact during operation.

Braking is a complex tribological phenomenon which occurs due to interaction of moving surfaces (i.e. brake block and steel rim). Brake block is a composite material which is employed to proper manveuring of the train. In braking, braking material (i.e. brake block) is pressed against the moving member by means of linkages or other means (like hydraulic or pneumatic). During braking, stationary member forces the moving member to stop by dissipating the kinetic energy (of moving member) into the atmosphere in the form of heat. During braking an extreme high temperature is reached at the shoe-tread interface which increases the chances of nucleation of surface cracks on the wheel and altering local and global stress-strain distribution. Braking effectiveness of shoe depends upon-
(i) wear rate (ii) coefficient of friction (iii) working temperature (iv) heat dissipating constant (v) application of constant pressure, etc.

***Corresponding Author,**
E-mail address: vishallife.97@gmail.com
All rights reserved: http://www.ijari.org

A desired brake must possess rigidity, flexible to accommodate, heat conducting, wear behavior (also wear of counterface material), environment friendly, economical etc.

From Designing to application of braking material, different parameters affects its efficiency like material used, binder resin, manufacturing process employed, resins, load, pressure (contact force), velocity, conditions of applications (temperature), attack angle, cooling medium used (wet), etching brake block, cast iron is used which produced noise during braking, in order to reduce rolling noise produced by the roughness of the tread wheel surface. To reduce the noise Fe-Cu-Cr-graphite sintered alloy is used. Investigation unfold that sintered alloy has constant and 1.8 times the coefficient of friction compared to the cast iron brake block moreover friction coefficient in cast iron shoe decreases while friction coefficient increases in sintered alloy with increase in sliding friction velocity[1-2] Investigated that phosphorous cast iron with different

with increase in sliding velocity rate of wear of cast iron increases up to a point then decreases significantly. With increase of both sliding velocity and contact pressure friction coefficient drops significantly at first, and then tends to be stable. Compacted graphite phosphorous cast iron has maximum coefficient of friction and minimum wear loss compared to grey phosphorous cast iron (both maximum wear rate and friction coefficient) and nodular phosphorous cast iron (good wear resistance but worst friction coefficient) [3]. Studied wear with silicon modified phenolic resin, straight phenol resin, boron phosphorus modified phenolic resin. Author investigation revealed variation of coefficient of friction with temperature due to thermal decomposition of binder resin. Friction materials with heat resistant pulp found to be with improved wear resistant and wear resistance is more pronounced at high temperature above the thermal decomposition temperature. At high temperature above critical temperature the thermal decomposition of resin of resin had a major impact on the abrasion rate, which contains random chain separation and carbonization.

During application of brake, wearing of brake block takes place but the amount of wear also depends upon friction coefficient, velocity and pressure. On further investigation it was found that with increase in pressure and velocity interface temperature also rises, which expedite the wearing of the brake block [4]. As the high temperature is reached, this causes damage to the wheel. The different method by which damage to the surface of the wheel takes place is ratcheting, surface crack nucleation and wear. The mechanism of substance

is created. When the applied brake liner is released, it also includes the steel substrate, which probably promotes surface breakage. In addition, the abrasive waste produced produces advanced abrasive wear. The cracks thus formed can move above if the deposited fluid pierce through the opening (cracks) and under pressure of load, which results into breaking into smaller pieces or fragments [5-6].

Braking can be assisted with liquids for dissipation of heat generated at the interface of contacting surface so that life of shoe can be increased, chances of abrasive wear, adhesion wear can reduced to the minimal but the results were opposite. It was found that the coefficient of friction between the steel wheel and brake shoe in the wet state shows a significant decrease compared to the dry low speed position and the rate of decline is around 30%.Friction properties at

wet condition [7]. The water film reconstructs an instantaneous velocity of about .30m/s for poor water flow and .45m/s for strong water flow. Cultivated water in the gap produces a part of surface hypertension which causes serious gas damage to the brake shoes (pitting wear/damage) .A powerful water flow leads to greater wear loss and lower interface temperature at high speed (6.70-13.40m/s) [7]. Also with increase in speed, attack angle and axle load resulted in the wheel/rail adhesion coefficient being significantly reduced. Water lubricants can have a significant impact on the adhesive behavior of the wheel/rail. The addition of water to the wheel/rail surface would significantly reduce the coefficient of adhesion. Under water conditions, the adhesion coefficient would decrease with an increase in velocity and angle of attack [8].

2. Experimental Methodology

The linear reciprocating tribometer is a floor standing model used for determining the frictional force and coefficient of friction to estimate the wear rate of a sample material. It is designed in such a way that it allows user to characterized Tribological performance of the material and lubricants in a wide range of operating conditions ranging from fretting wear to large amplitude linear sliding conditions. It can accommodate a variety of sample geometry to create point, line and area contact. It is connected with pc for sequence control and data acquisition with software. The linear movement is provided by the oscillatory movement of servo motor which is mounted vertically below base plate with its shaft projecting above it, also a crank is fastened to the shaft. The reciprocating rod is a long shaft having a specimen holding arm at middle supported; the upper specimen pin is clamped between jaws of holder. The bottom specimen 50×50×5mm is seated with in a holder; the holder is in the form of shallow bath having capacity to hold 20ml of oil. The specimen is clamped above a flexure with a heater block below it; heater block has a hole for temperature sensor used to measure temperature of lower specimen. After this the load is applied by dead weights on loading pan which pulls the shaft down to press the upper specimen against lower specimen.

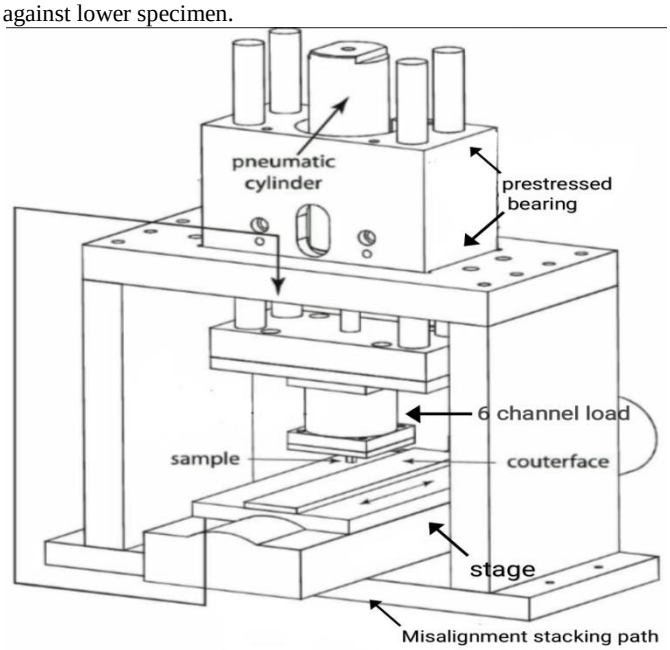


Fig. 1: Experimental test bench

The frequency of linear reciprocating tribometer may vary from 1Hz TO 50Hz depending upon the stroke length which also varies from 1mm to 20mm.the normal load can vary from 5N TO 50N in steps of 5N.the piezo sensor is used to measure the frictional force with a least count of 0.01N.

Author performed the experiment as shown in the figure. The experiments are held on the linear reciprocating tribometer to find the coefficient of friction and wear rate of the railway brake shoes at different temperatures. For Performing the experiment author prepared a piece of 25mm×25mm from railway brake block and placed it into the cabinet of the size 30mm. Authors also used a stainless steel pin having same composition to the railway wheel of 15mm stroke length holds at upper space of tribometer. The load is applied by upper space to the horizontal flat specimen. Here the load of 30N is applied at 15mm stroke length having frequency 10Hz.For first specimen we set the ambient temperature and time duration in the software and started the machine, upper specimen moves at a fixed stroke length on the specimen surface. The frictional force are measured by the sensors and displayed on the screen. After test the dimensional changes for both ball and specimen are used to measure the wear rate. Here we measure the weight of specimen before and after the test, there are some differences in their weight which we weigh by using the digital weighing machine having least count of 0.0001g and use to find the wear rate of brake shoes.

A similar phenomenon is used for the second specimen but with only difference is that the author set the temperature 45 degrees with the help of activating the inbuilt heater at the same load , stroke length and frequency. After the test we again weigh our piece as described above and determine its wear rate. We obtain some data in both the cases so we put these data in the software to get the graph and try to compare it.

Fig. 2: Brake shoe surface before test

Fig.3: Brake shoe surface after test

3. Result and Discussion

As earlier author specified that they take two specimen for the test, weight of first specimen is 10.0022g while weight of second specimen is 10.5522g and after the test on the linear reciprocating tribometer for fixed time of 180 seconds their weight become 10.0012g and 10.5494g respectively. So by this the wear of the brake shoe first specimen is 0.0010g while other is 0.0028g.

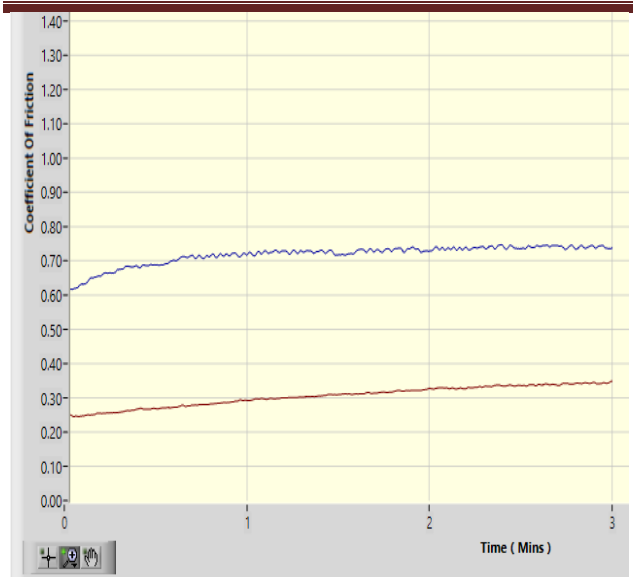


Fig.4: Coefficient of friction v/s time

As per the graph (i), the following points were noted as follows:-

1. In case of first specimen(red) which is tested at room temperature coefficient of friction varies in range of 0.25 to 0.35.While in case of second specimen(blue) which is tested at the 45 degree Celsius, the coefficient of friction varies from the range 0.61 to 0.74 which is much higher as compared to the first specimen.
2. During 0-1 min, in specimen first the coefficient of friction significantly increases from 0.25 to 0.30 with a difference of 0.05 while in specimen second here too the coefficient of friction rapidly increases from 0.61 to 0.72 with a difference of 0.11 which is again much higher as compared to specimen first.
3. During elapsed time 1min to 2min, In first specimen there is no appreciable change in coefficient of friction as slope is much less dominating (0min to 1min) during this coefficient of friction varies from 0.30 to 0.32. While in case of specimen second, the coefficient of friction varies from 0.72 to 0.735 and again the slope is not as dominating as in the case of 0-1 min.
4. During elapsed time 2min to 3min, as in first specimen the coefficient of friction varies from 0.32 to 0.35 and has an observable slope as compared to 1min to 2min. While in case of specimen second coefficient of friction is not so much very and its slope look like constant, during this the coefficient of friction varies from 0.735 to 0.74 with a difference of 0.005 which is lower as compared to specimen first on their difference.

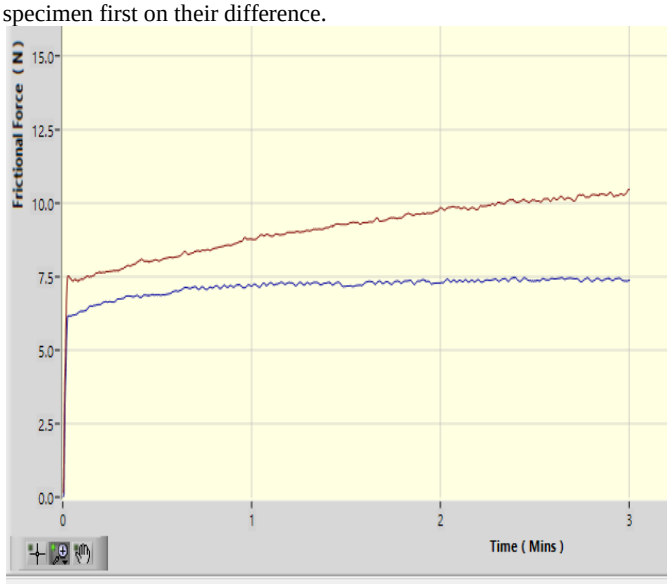


Fig.5: Frictional force vs time

As per the graph (ii), the following points were noted as follows:-

Frictional force directly depends upon (i) coefficient of friction and (ii) normal force in case of dry friction. There are two types of frictional force (i) static friction and (ii) kinetic friction. Static friction is the resistance offered by the object before coming into motion and kinetic friction is the resistance offered during motion. Static friction is always greater than kinetic friction. Static friction force can be calculated as the product of coefficient of friction and normal force acting on the body.

From the graph, frictional force vs time for two specimen shows that kinetic friction is less than static friction in both the cases. Moreover, it is evident that static friction is more at lower temperature than at high temperature. In specimen 1(red) static friction is more even though the coefficient of friction (0.25) is less it means that static friction is more at lower temperature and the rate at which friction force increases during 0-1 minute is comparatively significant and large as compared to others time elapse and in case of high temperature static friction is less even though the coefficient of friction is more at high temperature and increases more rapidly than at ambient temperature. Friction force nearly becomes constant after some time as their is no appreciable change in friction force during 1-3 minute. As wear rate is more in 2nd specimen it shows that high temperature influence the wear rate, even though the frictional force is greater at lower temperature wear rate is also low.

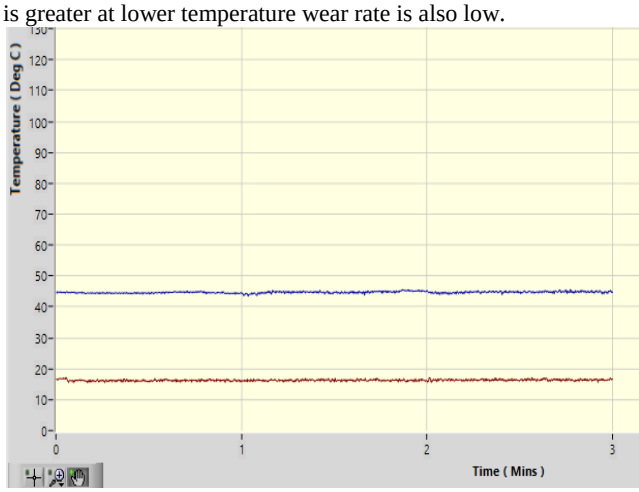


Fig.6: Temperature vs Time

From above graph, it is clear that there is no appreciable changes in temperature of the cavity on which thermocouple were embedded to measure the temperature. Test pieces after proper sizing were seated in the cavity, if during experiment temperature rises due to rubbing between test piece and pin, that would have caused some fluctuation in the temperature which would be recorded by the embedded thermocouple but co-generated during the experiment at interface was efficiently distributed and diffused in the atmosphere and hence there was no appreciable damage to piece which increases the operational life of the brake block.

4. Conclusions

From above discussion we found the following results are as follows:-

1. The coefficient of friction is much higher as compared to coefficient of friction at lab temperature.
2.It is found that coefficient of friction is found to start with high initial value and rise rapidly during 0-1min but in the rest of the time span there is no appreciable rate of change in friction coefficient with time in case of high temperature application while in case of lab temperature , coefficient of friction starts initially with lower value and progresses with nearly high constant rate during 0-1min as compared to the rest of time and attains a final value which is much lesser than the initial value at high temperature.

3. From graph 2, it is clear that frictional force is greater at low temperature as compared to high temperature even though friction coefficient is low for same parameter and maximum value attained at high temperature is lower than minimum value during low temperature. Both static and kinetic friction is greater in case of low temperature as compared to other, that means braking are more efficient at lower temperature.

4.The weight of pin remains unchanged during testing of 1st specimen and only the weight of the specimen reduces(i.e 0.0010g wear) it means that at lab temperature(18 degree celsius) wear of the wheel will be minimum which will increase the life of the wheel. But at high temperature there is wear in both pin and specimen (i.e wear of 0.0028g in specimen and 0.0002g in pin) which means at higher temperature wheel will wear along with the brake block and hence both the materials will wear resulting in net increase in operational cost of railway during operation at high temperature.

References

[1] C Ferrer, M Pascual, D Busquets, E Rayon. Tribological study of Fe-Cu-Cr graphite alloy and cast iron railway brake shoes by pin-on-disc technique, Wear 268, 2010, 784 789.

[2] Y Zhang, Y Chen, R He, B Shen. Investigation of Tribological properties of brake shoe materials phosphorous cast irons with different graphite morphologies, Wear, 166, 1993, 179 186

[3] US Hong, SL Jung, KH Cho, MH Cho, SJ Kim, H Jang. Wear mechanism of multiphase frictional materials with different phenolic resin matrices, Wear 266, 2009, 739 744.

[4] WJ Wang, F Wang, KK Gu, HH Ding, HY Wang, J Guo, QY Liu, MH Zhu. Investigation on braking tribological properties of metro brake shoe materials, Wear, 330-331, 2015, 498 506

[5] A Mazzu, L Provezza, N Zani, C petrogalli, A Ghidni, M Ficcoli. Effect of shoe braking on wear and fatigue damage of various railway wheel steel for high speed application, Wear 434 435, 2019, 203005.

[6] M Faccoli, L Provezza, A Mazzu, C Petrogalli, A Ghidni. Effect of full stop on shoe braked railway wheel wear damage, Wear, 428 429, 2019, 64 75

[7] F Wang, KK Gu, WJ Wang, QY Liu, MH Zhu. Study on braking tribological behaviors of brake shoe material under the wet condition, Wear, J.Wear, 9, 2015, 3

[8] WJ Wang, P Shen, JH Song, J Guo, QY Liu, XS Jin. Experimental study on adhesion behavior of wheel/rail under dry and water conditions, Wear 271, 2011, 2699 2705.

[9] L Cantone, E Crescentini, R Verzicco, V Vullo. A numerical model for the analysis of unsteady train braking and releasing manoeuvres, Proc. IMechE F: J. Rail Rapid Transit. 223, 2009, 305 317.

[10] A Pinchuk, P Dadoenkov, D Halutin, I Korago. Improved rolling system of railway stock on brake shoe, Procedia, 149, 2019, 258 263.

[11] P Wasilewski. Experimental study on the effect of formulation modification on the properties of organic composite brake shoes, J.Wear.8, 2017, 7

[12] MRK Vakkalagadda, DK Srivastava, A Mishra, V Racherla. Performance analyses of brake blocks used by Indian Railways, Wear, 328-329, 2015, 64 76.

Fish Physiol Biochem
https://doi.org/10.1007/s10695-020-00896-4

Effect of temperature on food consumption, immune system, antioxidant enzymes, and heat shock protein 70 of *Channa punctata* (Bloch, 1793)

Samar Pal Singh · Tauqueer Ahmad · JaiGopal Sharma · Rina Chakrabarti

Received: 20 April 2020 / Accepted: 3 November 2020
© Springer Nature B.V. 2020

Abstract The impact of water temperature on the physiology of *Channa punctata* (Bloch, 1793) was evaluated in the present study. Fish were acclimated at 25 ± 1 °C and then exposed at six different temperatures: 10, 15, 20, 25, 30, and 35 °C. *C. punctata* exposed at 10, 15, and 20 °C showed 30, 21, and 11% reduced food consumption, respectively compared to 25 °C. Significantly higher respiratory burst and myeloperoxidase activities were found in fish exposed at 20 and 25 °C after 12 h of exposure compared to other treatments. Nitric oxide synthase was significantly higher at 25 °C after 12 h and at 25 and 30 °C exposed fish after 7 days compared to others. The reduced glutathione level was significantly higher at 25 °C compared to other treatments after 7 days of exposure. The thiobarbituric acid reactive substances level was minimum at 25 °C. Significantly lower antioxidant enzymes, catalase, glutathione peroxidase, and glutathione S-transferase were found in gills of fish exposed at 25 °C compared to others in both samples. The highest antioxidant enzyme levels were found at 10 °C. Heat shock protein (Hsp) 70 levels were significantly lower in liver and muscle of fish exposed at 25 °C compared to other treatments. The Hsp level was significantly higher at 35 and 30 °C exposed fish compared to others after 12 h, and the level reduced after 7 days in these treatments. Thermal stress affects food consumption rate, immune system, antioxidant enzymes, and enzyme systems in fish. The elevated Hsp70 level serves as a biomarker of stress in *C. punctata*.

Keywords *Channa punctata* · Temperature · Food consumption · Myeloperoxidase · Antioxidant enzymes · Heat shock protein 70

S. P. Singh · T. Ahmad · R. Chakrabarti (✉)
Aqua Research Lab, Department of Zoology, University of Delhi, Delhi 110007, India
e-mail: aquaresearchlab@yahoo.co.in

J. Sharma
Department of Biotechnology, Delhi Technological University, Bawana Road, Delhi 110042, India

Introduction

The abiotic environmental factor temperature plays a significant role in the life of poikilotherm organism fish. Temperature is considered as one of the major environmental factors in aquatic ecosystem as it influences the growth and survival rate of fish (Brander 1994; Killen 2014; Sharma et al. 2014). Exposure of fish to adverse environmental temperature deleteriously affects their biological functions that may cause a series of biochemical and physiological changes. Seasonal and diurnal fluctuations of water temperature influence individual fish throughout its life (Basu et al. 2002). Temperature beyond the optimal limit of a particular species has an adverse impact on fish health by modulating the metabolic rate, oxygen consumption, and the virulence of pathogens that in turn may cause a variety of pathophysiological disturbances (Gordon 2005; Dalvi et al. 2009). The food consumption rate and digestive enzyme activities are indicators of health status of fish and exposure

Published online: 19 November 2020

of fish beyond their thermal tolerance limits affects these parameters (Chakrabarti and Sharma 2005). A shift in temperature range from the optimum changes the enzymatic activities. The food consumption rate and digestive enzyme activities in catla *Catla catla* and magur *Clarias batrachus* were reduced at < 25 °C (Singh et al. 2013; Sharma et al. 2014). Reduced food consumption rate makes the fish weak and prone to the diseases.

The rapid change of water temperature produces stress in fish and thereby affects the fish immune system (Mariana et al. 2019). The innate immune system is the fundamental defense mechanism in fish. Some of the immunological parameters, *viz*., respiratory burst activity, myeloperoxidase, and nitric oxide synthase, serve as bioindicators. The lysosomal protein myeloperoxidase is stored in the neutrophils, and it produces reactive oxygen species (ROS) like hypohalous acid (HOCl, HOBr, and HOI) and other toxic oxidants from hydrogen peroxide and halides. These oxidants oxidize organic molecules and have antimicrobial function (Kinkade Jr et al. 1983; Klebanoff 2005; Davies 2011). The nitric oxide synthase helps in the production of nitric oxide, the signaling molecule from L-arginine (Andrew and Mayer 1999). The inducible nitric oxide synthase also produces nitric oxide that shows a direct anti-microbial activity (Schairer et al. 2012).

The enhanced temperature results in higher oxygen consumption. Therefore, may up-regulate reactive oxygen species production as a side products of enhanced metabolism (Lushchak 2011). The imbalance between the production of oxidant and antioxidants components results into oxidative stress. Many of these compounds play an important role in metabolic pathways, essential for the defense system of the organism, and their byproducts show bactericidal activities (Billar and Takashi 2018). The antioxidant enzymes (*viz*., glutathione peroxidase, catalase, superoxide dismutase) stabilize and deactivate free radicals before they attack different components of cell (Blokhina et al. 2003; Noori 2012). Glutathione peroxidase protects the cells from free radicals, hydrogen, and lipid peroxides (Kohen and Nyska 2002), and glutathione S-transferases are crucial detoxification enzymes which mainly present in cytosol (Coles and Kadlubar 2003). Antioxidant enzymes and lipid peroxidation are the potential biomarkers of environmental temperature exposure in different organisms (Verlecar et al. 2007; Bocchetti et al. 2008). In the cell membrane, the reactive oxygen species causes lipid peroxidation and produces lipid hydroperoxide (LOOH) that further breaks down into aldehydes (malonaldehyde, isoprotans, hydrocarbons etc.). This affects the integrity of the membrane (Beevi et al. 2010). Heat shock proteins (Hsp), also known as stress proteins, are key markers to assess environmental stress (Iwama et al. 1999, 2004). Among various Hsps, Hsp70 is the most abundant and serves as an indicator of thermal stress (Nie et al. 2017; Saranyan et al. 2017). It plays a significant role in the tissue immune response (Fu et al. 2011).

Spotted snakehead *Channa punctata*, Bloch (Channidae family, Perciformes order) is an economically important freshwater, carnivore fish of South-East Asian countries (Talwar and Jhingran 1991; Froese and Pauly 2019). Usually, the fish is found in the derelict, swampy water bodies with vegetation. The air-breathing capacity makes the fish hardy and helps it to withstand low dissolve oxygen level of the water. In a natural water body, it breeds in the monsoon months and shows parental care. *C. punctata* is found around the temperature of 22–28 °C (Froese and Pauly 2019). The fish is marketed in a living condition; it has consumers' preference for taste and nutritional and medicinal values (Haniffa et al. 2004). It is recommended in diet during convalescence (Marimuthu et al. 2009). So far, *C. punctata* is harvested from the natural water bodies; gradually, their number is decreasing due to habitat destruction and over fishing (Hossen et al. 2017). Therefore, development of proper husbandry is essential for the sustainable production and supply of this nutritious food fish. *C. punctata* is a good candidate for intensive culture as it tolerates a low dissolved oxygen level. Therefore, the impact of the abiotic factor, temperature on the food consumption rate, and immunological and stress biomarkers should be evaluated for the development of the proper culture technique of *C. punctata*. These biomarkers will help to understand the optimum culture conditions for this fish.

Materials and methods

Experimental conditions

Channa punctata (90 fish, length 18.2 ± 1.0 cm; wet weight 65.3 ± 5.42 g) were collected and transported from a local fish market and randomly distributed in glass aquaria (60 L each) maintained in the wet laboratory facility. The stocking density of fish was five fish aquarium^{-1}. The fish were acclimated at 25 ± 1 °C for 21

days to avoid the transportation stress. Then, the fish were exposed at six different temperatures of 10, 15, 20, 25, 30, and 35 °C. In the present study, the 10–35 °C temperature was selected based on the range of temperature found in India in different months (Ahmad et al. 2014). As fish were acclimated at 25 °C prior to exposure, this was considered a control temperature. Three replicates were used for each temperature. The photoperiod was maintained as 12 h:12 h light and dark. Fish were fed with laboratory made pelleted feed at the rate of 5% of body weight throughout the experiment. The feed was composed of dry fish powder, wheat flour, cod liver oil, and vitamin-mineral premix. The protein and lipid contents of prepared feed were 42 and 10%, respectively.

Each aquarium (60 L) was connected with one filtration unit (Sera fil bioactive 130, Germany) and one cooling unit (Julabo F34, Julabo Labortechnik GmbH, Germany) or heating (Sera Aquarium Heater 300, Germany) unit for the maintenance of desirable temperature. A thermostatically controlled heater (Hicon, India) was used for the maintenance of water temperature at 35 °C. The water from fish culture aquarium was first circulated into the filtration unit, then to the cooling unit, and finally to the culture aquarium. The acclimation temperature was 25 °C, and the water temperature was either increased or decreased to achieve the assigned experimental temperature. The rate of temperature change was 1 °C per 12 h (i.e., 2 °C day^{-1}). This required 2.5–7.5 days to reach the assigned temperature. The exact time of reaching the assigned temperature was recorded. The first sample was collected after 12 h and the second one after 7 days of reaching the assigned temperature. The total experimental exposure period was 7 days after reaching the assigned temperature. Water quality parameters, *viz*., temperature, pH, and dissolved oxygen, were monitored with HQ40d Multiparameter (Hach, USA) using a specific electrode. Dissolved oxygen level was maintained at 5.0 mg L^{-1} in all the treatment with the help of an aerator. The pH of water ranged from 7.7 - 8.0 during the study period.

The whole study was conducted following the guidelines of Institutional Animal Ethics Committee, IAEC (approved by Committee for the Purpose of Control and Supervision of Experiments on Animals, CPCSEA).

Survival rate and food consumption

Survival rate was monitored regularly. During each feeding, fish were allowed to feed for 1 h. Then unconsumed food was siphoned from individual aquarium and oven dried, and the weight was recorded. The difference between the mass of the food given and the remaining food in the aquarium after 1 h was considered as food consumption rate of individual fish and expressed as g $fish^{-1}$.

Blood and tissue collection

Six fish (2 fish per aquarium × 3 replicates) per treatment (in each sampling) were collected after 12 h and 7 days of reaching the assigned temperature and anaesthetized with tricaine methanesulfonate, MS-222 (Sigma-Aldrich, USA) without showing any adverse effect on fish (Palić et al. 2006). Blood was collected from caudal vein of individual fish using 2 mL of disposable syringe (0.55 × 25 mm/24 × 1). Two different aliquots of blood were used for different analyses. The first aliquot was transferred to a micro-centrifuge tube (1.5 mL) coated with 2.7% ethylenediamine tetraacetic acid (EDTA) for respiratory burst activity estimation. The second aliquot was transferred to a micro-centrifuge tube (without EDTA), allowed to clot at room temperature for 1 h, stored at 4 °C, and centrifuged at 1500×*g* for 10 min. Then the serum was stored at − 20 °C and used for assays. Muscle, gills, and liver tissues were collected aseptically, frozen in liquid nitrogen, and stored at − 80 °C.

Biochemical assays

Respiratory burst activity

The respiratory burst activity was assayed using nitroblue tetrazolium, NBT (Anderson and Siwicki 1995). In 20 μL of blood sample, 20 μL of nitroblue tetrazolium (0.2%, Himedia, India) was added and kept at 25 °C for 30 min. In a separate tube, 25 μL of this mixture was taken, and dimethyl formamide (500 μL; Merck, India) was added to this mixture. This dissolved the formazan produced from nitroblue tetrazolium. The whole mixture was centrifuged at 2000×*g* for 5 min; the supernatant was collected, and the absorbance was recorded at 540 nm using a microplate reader (BioTek Synergy ^{TM} H1, USA).

Myeloperoxidase activity

The myeloperoxidase activity of serum was measured using a Hank's balanced salt solution (HBSS) without Ca^{2+}, Mg^{2+}, and phenol red (Quade and Roth 1997). In an individual well of a 96-well microplate, 90 μL of HBSS (Himedia) and 10 μL of serum were added. Then, a 35 μL solution of 20-mM 3,3',5,5'-tetramethylbenzidine hydrochloride and 5 mM H_2O_2 (Genei, India) were added into each well and incubated for 2 min, and 4 M sulfuric acid (35 μL) was added to stop the reaction. Optical density was measured at 450 nm.

Nitric oxide synthase

Nitric oxide synthase was assayed in muscle (Lee et al. 2003). In a 1-mL phosphate buffer (pH 7.4), a 100-mg tissue was homogenized and centrifuged at 10,500×*g* for 20 min at 4 °C. The supernatant (100 μL) was mixed with a 100 μL Griess reagent and incubated at 25 °C for 10 min. The Griess reagent was composed of 1% sulfanilamide, 0.1% N-(1-naphthyl) ethylenediamine, and 5% phosphoric acid (Himedia). The absorbance was recorded at 540 nm. The nitrite standard curve was prepared, and the concentration was expressed as mmol mg^{-1} tissue (wet weight basis).

Thiobarbituric acid reactive substances

In the muscle, thiobarbituric acid reactive substances (TBARS) was assayed (Ohkawa et al. 1979). In a 1.15% KCl solution (450 μL), the muscle (50 mg) was homogenized, and the sample was incubated for 1 h at 100 °C in acid medium (containing 0.45% sodium dodecyl sulfate and 0.6% thiobarbituric acid). After cooling, the sample was centrifuged at 825×*g*, and the absorbance of the supernatant was measured at 532 nm. The standard curve was prepared with 1,1,3,3-tetramethoxy propane (Himedia), and the concentration was expressed as μmol malondialdehyde mg^{-1} protein.

Reduced glutathione

Reduced glutathione (GSH) activity of gills was determined (Jollow et al. 1974). In a 10 mL phosphate buffer (0.1 M, pH 7.4), the gill tissue (1 g) was homogenized and centrifuged at 10,500×*g* for 20 min; the supernatant was treated as post-mitochondrial supernatant (PMS). The supernatant was precipitated with 4% sulfosalicylic acid (Himedia), kept at 4 °C for 1 h, and then centrifuged at 1500×*g* for 15 min at 4 °C. This supernatant was used for GSH assay. The assay mixture consisted of supernatant, 0.1 M phosphate buffer (pH 7.4), and 5-5-dithiobis-2-nitrobenzoic acid (DTNB, Himedia). Optical density was measured at 412 nm and expressed as μmol of DTNB g^{-1} tissue (wet weight basis).

Catalase

Catalase activity of gills was assayed using the method of Claiborne (1985). In a 10 μL of sample, 140 μL phosphate buffer (0.1 M, pH 7.4) was added and mixed well. The reaction was initiated with the addition of 50 μL of hydrogen peroxide (H_2O_2, 0.02 M). The change in absorbance was recorded at 240 nm. The activity was expressed as nmol of H_2O_2 consumed mg^{-1} protein min^{-1}.

Glutathione peroxidase

The glutathione peroxidase (GPx) activity was estimated in gills (Mohandas et al. 1984). The composition of the reaction mixture was as follows: 140 μL phosphate buffer (0.1 M, pH 7.4), 10 μL EDTA (1 mM), 10 μL sodium azide (1 mM, CDH, India), 10 μL glutathione reductase (1 IU mL^{-1}, Sigma, USA), 10 μL reduced glutathione (GSH, 1 mM, Sigma), 10 μL nicotinamide adenine dinucleotide phosphate reduced (NADPH, 0.2 mM, Himedia), 10 μL hydrogen peroxide (0.25 mM, Merck, Germany), and 10 μL sample. Oxidation of NADPH was recorded at 340 nm. The enzyme activity was calculated and expressed as nmol of NADPH oxidized mg^{-1} protein min^{-1} using a molar extinction coefficient of 6.22×10^3 M^{-1} cm^{-1}.

Glutathione S-transferase

Glutathione S-transferase (GST) activity of gills was assayed (Habig et al. 1974). The composition of the reaction mixture was as follows: 30 μL supernatant, 147.5 μL phosphate buffer (0.1 M, pH 7.4), 2.5 μL 1-chloro-2-dinitrobenzene (CDNM, 0.60 mM, Sigma), and 20 μL reduced glutathione (10 mM, Himedia). Absorbance was monitored at 340 nm in kinetic mode with a BioTek microplate reader using Gen 5 software. The software helps to correct the path length automatically. Activity was expressed as nmol of CDNB mg^{-1} protein min^{-1}. The molar extinction coefficient is 9.6×10^3 M^{-1} cm^{-1}.

Heat shock protein 70

Fish Hsp70 ELISA kit (CUSABIO, Wuhan, China) was used for the estimation of heat shock protein (Hsp) 70 of muscle and liver. Tissue (100 mg) was rinsed with × 1 phosphate buffer saline (pH 7.4) and homogenized in 1 mL PBS using RQ-129/D (REMI, India) and stored overnight at − 20 °C. After two times of freeze-thawing (following the protocol of the manufacturer), the homogenate was centrifuged at 5000×*g* for 5 min at 4 °C. The supernatant was stored at − 80 °C for assay. Sample (50 μL) was added to the ELISA plate (96 wells plate) that was pre-coated with antibody. Conjugate (50 μL) was added to each well, except the blank; mixed well; and incubated at 37 °C for 60 min. Each well was aspirated and washed with 200 μL of washing buffer three times. A 50 μL of horseradish peroxidase-avidin (HRP-avidin) was added to each well, except the blank, and mixed properly and then incubated for 30 min at 37 °C. Aspiration was repeated for three times. Substrates A and B (each 50 μL, provided with the kit) were added to each well and incubated at 37 °C for 15 min. Stop solution (50 μL) was gently added to this and mixed properly. Absorbance was recorded at 450 nm, and concentration of Hsp70 was expressed (following the protocol of manufacturer) in pg mL^{-1}(El-Ansary et al. 2012; Singh et al. 2015).

Total protein

The protein content of each tissue was measured using Folin Ciocalteus phenol reagent (Lowry et al. 1951). The optical density was measured at 750 nm. The standard was prepared with bovine serum albumin (BSA).

Statistical analysis

Data were compiled as mean ± standard error. The regression analysis showed the relationship between the temperature and the food consumption rate of fish. The Shapiro-Wilk test (1965) showed the normality of different parameters (data). The multivariate test two-factor MANOVA was performed, and the Wilks' Lambda row showed that the two-way MANOVA was statistically significant ($P < 0.001$). The parameters are significantly dependent on time and temperature ($P < 0.001$). Further analysis was performed to determine how the dependent variables differ among the different levels of temperature and time. The Tukey's test (2013) showed the significant differences among the treatments. The SPSS software 26.0 (South Asia (P) Ltd., Bangalore, India) was used for analysis. Statistical significance was accepted at $P < 0.05$ level.

Results

Survival rate and food consumption

All fish survived in each culture unit. There was no mortality of fish. The results of food consumption rate are shown in Fig. 1. The food consumption rate showed direct relationship with water temperature. *C. punctata* consumed less food as the water temperature reduced from 25 °C. In the 10 °C treatment, as the temperature dropped to 15 °C, fish consumed 14.5% less food, and at 10 °C, 30% reduced food consumption was recorded. It required 7.5 days to reach 10 °C from the acclimation temperature 25 °C. The food consumption rate of *C. punctata* was recorded on day 8 (one day) of reaching the assigned temperature. A sixth-degree polynomial relationship ($R^2 = 0.81$) was found between the water temperature and the food consumption rate of fish (Fig. 1a). In the 15 °C treatment, when temperature dropped to 20 °C, 10% less food consumption was recorded, and it was reduced 21% at 15 °C. On day 6, 15 °C was reached, and the food consumption rate was recorded for 4 days after reaching the assigned temperature. A fourth-degree polynomial relationship ($R^2 = 0.81$) was found between the water temperature and the food consumption rate of fish (Fig. 1b). It required 2.5 days to reach 20 °C from 25 °C, and food consumption rate was recorded for consecutive 6 days after reaching the assigned temperature. In the 20 °C treatment, 11% reduced food consumption was recorded on the first day of reaching this temperature compared to the consumption rate at 25 °C. After 6 days of exposure, the food consumption rate further reduced to 24% compared to 25 °C treatment. Food consumption rate showed a sixth-degree polynomial relationship with the decreasing water temperature ($R^2 = 0.99$) in this treatment (Fig. 1c). There was no significant difference in food consumption rate of fish exposed at 30 and 35 °C compared to the fish at 25 °C.

Biochemical assays

Respiratory burst activity

Respiratory burst activity was significantly higher in fish exposed at 20 and 25 °C compared to other treatments after 12 h. After 7 days, highest activity was found at 25 °C and this treatment was followed by 15 and 30 °C. The lowest activity was found in fish exposed at 35 and 10 °C after 12 h and 7 days, respectively (Table 1).

Myeloperoxidase activity

Myeloperoxidase activity was significantly higher in fish exposed at 20 and 25 °C compared to other treatments after 12 h. There was no significant difference between these treatments. After 7 days of exposure, myeloperoxidase activity was significantly higher in fish exposed at 25–35 °C compared to fish exposed at 10–20 °C (Table 1).

Nitric oxide synthase

Nitric oxide synthase level was significantly higher in muscle of fish exposed at 25 °C after 12 h and 25 and 30 °C after 7 days of exposure compared to other treatments. Nitric oxide synthase level was minimum at 10 °C treatment in both samples (Table 1).

Thiobarbituric acid reactive substances

Thiobarbituric acid reactive substances (TBARS) level was significantly higher in fish exposed at 10 °C compared to other treatments after 12 h of exposure. After 7 days, TBARS level increased in fish exposed at 30 and 35 °C compared to the 12 h exposed fish of the same treatment. The TBARS level was minimum in fish cultured at 25 °C in both days of sampling (Table 1).

Reduced glutathione

The reduced glutathione (GSH) level was significantly higher in gills of fish exposed at 25 °C compared to other treatments after 12 h. A similar trend was found after 7 days of exposure. There was no significant difference in GSH level among fish exposed at 10, 15, and 35 °C after 12 h and 7 days of exposure (Table 1).

Catalase

In gills, significantly higher catalase activity was found at 10 °C compared to other treatments after 12 h of exposure. This group was followed by fish exposed at 35 °C. After 7 days, the catalase level increased significantly in 10–20 °C treatments compared to the activity found after 12 h of exposure in the respective treatment, whereas the catalase activity reduced in 30–35 °C treatments after 7 days compared to the 12 h exposed fish in the respective treatment. Catalase level was always minimum in fish exposed at 25 °C (Table 1).

Glutathione peroxidase

The effect of temperature change was recorded on glutathione peroxidase (GPx) level of gills. GPx level increased when the water temperature decreased from 25 °C to 20–10 °C after 12 h of exposure. It was interesting to record that GPx level was also higher in fish exposed at 30–35 °C compared to the fish exposed at 25 °C. Significantly lower GPx level was found in fish exposed at 25 °C in both samples. GPx level reduced significantly in all treatments after 7 days compared to the 12 h exposed fish in the same treatment, except the 20 °C treatment (Table 1).

Glutathione S-transferase

Significantly, the lower level of glutathione S-transferase (GST) was found in gills of fish exposed at 25 °C compared to other treatments in both samples. In gills of 10 °C exposed fish, 4- and 4.65-fold higher GST levels were found compared to control fish after 12 h and 7 days, respectively (Table 1).

Heat shock protein70

In the liver of *C. punctata*, Hsp70 level was significantly higher at 35 °C after 12 h of exposure. This group was followed by 30 °C treatment. Hsp70 was always minimum in fish exposed at 25 °C. After 7 days of exposure, Hsp70 level reduced significantly in all treatments compared to the 12 h exposed fish of same temperature (Fig. 2a). Like the liver, significantly higher level of Hsp70 was found in the muscle of fish exposed at 30 and 35 °C compared to the other treatments after 12 h of exposure. After 7 days, Hsp levels reduced in fish exposed at 30 and 35 °C compared to the 12 h exposed fish in the same

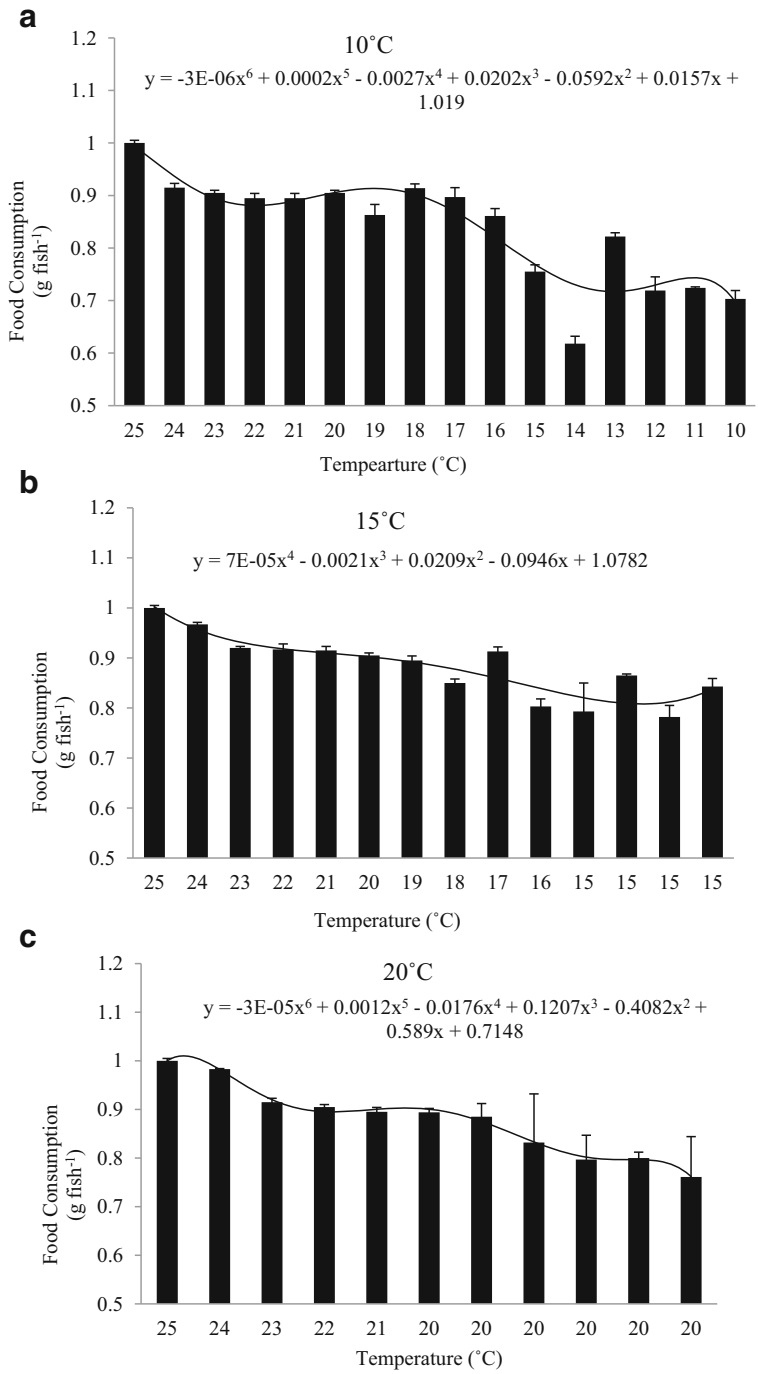


Fig. 1 Effect of temperature on the food consumption rate of *Channa punctata* exposed at (**a**) 10, (**b**) 15, and (**c**) 20 °C temperatures. Water temperature gradually decreased from acclimation temperature of 25 to experimental temperatures 10, 15, and 20 °C. It required 2.5, 5, and 7.5 days to reach 10, 15, and 20 °C, respectively, from the acclimation temperature 25 °C. Food consumption rate was recorded for 6, 4, and 1 day(s) in 20, 15, and 10 °C treatments, respectively. Bar (means ± SEM) indicates the food consumption of fish at various temperatures (3 replicates)

treatment. Hsp70 level was minimum in fish exposed at 25 °C in both samples (Fig. 2b).

Discussion

Water temperature is an important factor that influences the physiology and survival rate of the aquatic species. Lowering of temperature from the acclimation temperature caused mortality of catla and magur (Singh et al.

Table 1 The respiratory burst activity, serum myeloperoxidase level, nitric oxide synthase, and thiobarbituric acid reactive substances (TBARS) levels of muscle and reduced glutathione, catalase, glutathione peroxidase, and glutathione-S-transferase levels of gills found in *C. punctata* exposed at different temperatures

	Sampling	10 °C	15 °C	20 °C	25 °C	30 °C	35 °C
Respiratory burst activity (O.D. at 540 nm)	12 h	0.72 ± 0.02^{b}	0.84 ± 0.03^{b}	0.96 ± 0.03^{a}	0.99 ± 0.06^{a}	0.81 ± 0.06^{b}	0.69 ± 0.07^{b}
	7 days	0.42 ± 0.03^{b}	0.52 ± 0.04^{b}	0.62 ± 0.07^{b}	0.88 ± 0.07^{a}	0.72 ± 0.02^{a}	0.74 ± 0.07^{a}
Myeloperoxidase (O.D. at 450 nm)	12 h	2.28 ± 0.05^{b}	2.52 ± 0.11^{b}	2.94 ± 0.11^{a}	2.85 ± 0.1^{a}	2.21 ± 0.10^{b}	2.18 ± 0.13^{b}
	7 days	2.16 ± 0.03^{c}	2.42 ± 0.04^{b}	2.51 ± 0.07^{b}	2.8 ± 0.05^{a}	2.70 ± 0.05^{a}	2.74 ± 0.06^{a}
Nitric oxide synthase (mmol mg^{-1} tissue)	12 h	6.3 ± 0.04^{d}	6.5 ± 0.05^{cd}	8.0 ± 0.71^{c}	18.2 ± 0.67^{a}	16.2 ± 0.52^{b}	9.2 ± 0.22^{c}
	7 days	6.7 ± 1.03^{c}	7.8 ± 0.6^{c}	10.4 ± 0.67^{bc}	18.3 ± 0.85^{a}	17.2 ± 0.51^{a}	11.5 ± 0.15^{b}
TBARS (μmol MDA mg^{-1} protein)	12 h	252 ± 18^{a}	211 ± 11^{b}	208 ± 6^{b}	167 ± 8^{c}	187 ± 10^{c}	200.0 ± 7^{b}
	7 days	214 ± 7^{a}	194 ± 6^{b}	190 ± 7^{b}	175 ± 5^{c}	209 ± 17^{a}	220 ± 11^{a}
Reduced glutathione (μmol DTNB g^{-1} tissue)	12 h	37 ± 1^{c}	36 ± 1^{c}	66 ± 2^{a}	73 ± 2^{a}	50 ± 3^{b}	37 ± 1^{c}
	7 days	33 ± 1^{c}	37 ± 1^{c}	50 ± 3^{b}	76 ± 1^{a}	49 ± 2^{b}	39 ± 2^{c}
Catalase (nmol H_2O_2 min^{-1} mg^{-1} protein)	12 h	254 ± 1^{a}	182 ± 1^{c}	120 ± 5^{d}	77 ± 2^{e}	110 ± 3^{d}	234 ± 1^{b}
	7 days	746 ± 14^{a}	372 ± 11^{b}	201 ± 11^{c}	75 ± 3^{e}	84 ± 3^{e}	154 ± 2^{d}
Glutathione peroxidase (μmol NADPH min^{-1} mg^{-1} protein)	12 h	1471 ± 101^{a}	1129 ± 61^{b}	604 ± 40^{d}	419 ± 24^{e}	563 ± 27^{d}	840 ± 21^{c}
	7 days	903 ± 38^{a}	637 ± 35^{b}	630 ± 32^{b}	222 ± 12^{d}	319 ± 27^{c}	683 ± 24^{b}
Glutathione-S-transferase (nmol CDNB min^{-1} mg^{-1} protein)	12 h	1942 ± 40^{a}	1433 ± 56^{b}	1196 ± 52^{c}	475 ± 25^{e}	924 ± 33^{d}	1183 ± 69^{c}
	7 days	2143 ± 50^{a}	1306 ± 4^{7b}	1048 ± 49^{c}	461 ± 21^{d}	1039 ± 55^{c}	996 ± 72^{c}

Values with different superscripts in same row are significantly ($P < 0.05$) different (2 fish per aquarium × 3 replicates)

2013; Sharma et al. 2014). This was very interesting that there was no mortality of *C. punctata* at 10 °C. In the present study, reduced food consumption was recorded with decreasing water temperature. The earlier study with catla and magur also showed that food consumption rates were affected at temperature below 25 °C (Ahmad et al. 2014; Sharma et al. 2017). Like catla and magur, food consumption rate of spotted snakehead was also unaffected at temperatures 30–35 °C. The low temperature has a significant negative impact on fish appetite, growth, and activity (Brown et al. 1989; Clark et al. 1995; Bendiksen et al. 2002). Fish are exposed at low temperature during winter that results into reduced food consumption and poor fish growth. Water temperature below 25 °C affected the food consumption rate and digestive enzyme activities in fish, and these parameters served as indicators of stress in fish (Ahmad et al. 2014).

In *C. punctata*, the impact of temperature change on the immune system was documented. A significant reduction in respiratory burst, myeloperoxidase, and nitric oxide synthase levels were observed in spotted snakehead exposed at < 25 °C, except similar levels of respiratory burst and myeloperoxidase were found in fish exposed at 20 and 25 °C after 12 h of exposure. These parameters are indicators of compromised immune system of fish. The respiratory burst activity is an important oxygen-dependent pathogen-killing mechanism in phagocytic cells, like monocytes, macrophages, and neutrophils. In fish, non-specific immune activity plays a significant role to give protection against pathogens (Haugland et al. 2012). Myeloperoxidase is released from cytoplasmic granules of neutrophils and monocytes. It reacts with the H_2O_2 produced in respiratory burst to form a complex which can oxidize a variety of substances. The hypochlorous acid, the product of this reaction, is cytotoxic, and it helps neutrophil to kill bacteria and other pathogens (Klebanoff 1999). The reduced values of respiratory burst activity and myeloperoxidase at 10–20 °C after 7 days compared to 12 h exposed *C. punctata* showed the long-term effect of temperature mediated stress in fish. A decreased level

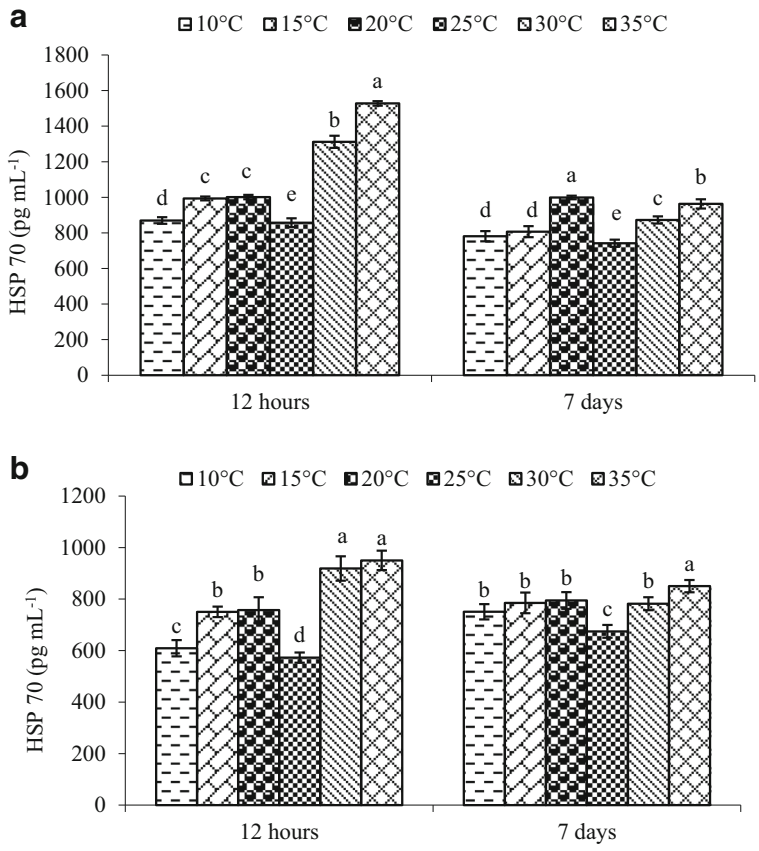


Fig. 2 Effect of temperature on the Hsp70 levels in (**a**) liver and (**b**) muscle of *Channa punctata* exposed at various temperatures. Bars (means ± SEM) with different superscripts are significantly different (2 fish per aquarium × 3 replicates)

of nitric oxide synthase was observed with the lowering of water temperature from the acclimation temperature, 25 °C. A similar result was reported in catla exposed at low temperature (Sharma et al. 2017). After 7 days of exposure, the elevated level of nitric oxide synthase was recorded compared to the 12 h exposed fish at 10–20 °C.

The exposure of *C. punctata* at 10–20 °C for 12 h caused a significant oxidation in the muscle as TBARS level was higher in these treatments compared to 25–30 °C treatments. The increased lipid peroxidation and disturbed antioxidant enzymes resulted in additional stress in spotted snakehead at low temperature. Significantly higher lipid peroxidation was found in spotted snakehead exposed at 32 °C compared to the control group kept at 20 °C (Kaur et al. 2005). A 2.7-fold increase in lipid peroxidation was observed in juvenile European seabass *Dicentrarchus labrax* during post-thermal stress (Vinagre et al. 2012). Like 10 °C exposed *C. punctata*, significantly higher TBARS levels were also found at 30–35 °C exposed fish (compared to 25 °C) after 7 days of exposure. The exposure of bald notothen *Pagothenia borchgrevinki* at 4 °C for 3 weeks caused higher levels of lipid peroxides (Almroth et al. 2015). This showed the impact of long duration exposure in the muscle.

GSH level was affected in gills of *C. punctata* exposed at low and high temperatures compared to the control group. A similar result was also observed in the same species exposed at 32 °C (Kaur et al. 2005). In temperature stress, probably thiols content of the cells is modulated to overcome the effect of reactive oxygen species. Therefore, they are the first to be used in cellular defense against oxidative stress (Lushchak and Bagnyukova 2006a, 2006b; Bagnyukova et al. 2007). In gills of spotted snakehead, antioxidant enzymes catalase, GPx, and GST levels were enhanced when the water temperature dropped from 25 to 20–10 °C and also increased from 25 to 30–35 °C. An increase in ambient temperature caused higher antioxidant enzyme activities. This may be related to the higher activity of glutathione-dependent enzymes, especially GPx and GST. These enzymes played an important role in antioxidant defenses against oxidative stress (Lushchak and Bagnyukova 2006a, 2006b; Bagnyukova et al. 2006, 2007). GPx catalyzes the reduction of H_2O_2 using reduced glutathione and provides protection to the animal cells against oxidative damage (Kohen and Nyska 2002). Higher antioxidant enzyme was found in European seabass exposed at different temperatures (Vinagre et al. 2012). The exposure of bald notothen *P. borchgrevinki* at 4 °C for 12 h showed enhanced levels of antioxidant enzymes, *viz*., glutathione reductase, glutathione peroxidase, and glutathione S-transferase levels (Almroth et al. 2015). After 7 days, catalase levels in gills of *C. punctata* showed different responses compared to the 12 h exposed fish in the same treatment. Catalase level increased at 10–20 °C, whereas the level reduced at 25–35 °C. In the present study, increased levels of antioxidant enzymes were also found at low temperature. The cold temperatures are accompanied by an elevated physical dissolution of oxygen and may enhance the availability of oxygen to ROS-generating processes at low temperatures (Grim et al. 2010). The two-factor MANOVA shows that the temperature has a statistically significant effect on all parameters ($P < 0.001$); the sampling time, i.e., first sampling (after 12 h of reaching the assigned temperature) and second sampling (after 7 days), also shows a significant impact on the parameters, except for the dependent variables myeloperoxidase, TBARS, and GST (Table 1). Further, significant interaction has been found between temperature and time for all parameters, except nitric oxide synthase.

Heat shock proteins are another protective mechanism of cellular system that protect cell against heat and cold shock and environmental contaminants. Various studies suggested that cellular stress response modulates the Hsps and enhances the survival and health of the stressed fish (Wang et al. 2007; Liu et al. 2017; Sales et al. 2019). Higher level of Hsp70 was found in liver and muscle of *C. punctata* exposed at 30–35 °C compared to the other treatments after 12 h of exposure. An acute and chronic temperature exposure of sea bream *Spondyliosoma cantharus* resulted into increased Hsp70 expression in fish exposed at a high temperature (32 °C) compared to fish maintained at 25 °C (Deane and Woo 2005). Increased levels of Hsp70 were observed in Florida pompano *Trachinotus carolinus* and spotted sea bass *Lateolabrax maculates* in acute thermal stress (Cardoso et al. 2015; Shin et al. 2018). In the present study, the increased Hsp70 level might be linked with an antioxidant enzyme. In *C. punctata*, elevated levels of Hsp70 and catalase were found at 30–35 °C after 12 h of exposure. After 7 days, the expressions of Hsp70 were lower in the liver and muscle, and also the catalase level reduced compared to 12 h exposed fish at 30 and 35 °C. The increased levels of Hsp70 and catalase were found in mud crab

Scylla paramamosain exposed at 35 °C (Liu et al. 2018). The elevated level of Hsp70 lasted for a few hours (Somero 2002). Tissue-specific expression of Hsp70 was also recorded in the present study as the higher level of Hsp70 was found in the liver compared to the muscle of fish exposed at the same treatment. In western painted turtle *Chrysemys picta*, around 3- and 4-fold higher expressions of Hsp 73 were found in the brain and liver compared to the skeletal muscle (Scott et al. 2003). In thermal stress, tissue-specific regulation of *HSF1* was found in zebrafish *Danio rerio* (Rabergh et al. 2000). The study showed the presence of two *zHSF1* mRNA forms that were expressed in a tissue-specific fashion in heat stress.

Conclusions

In *Channa punctata*, temperature below 25 °C affected food consumption rate, immune system, antioxidant enzymes, enzyme systems, and Hsp70 and caused lipid peroxidation. Temperature above 25 °C affected antioxidant enzymes and enzyme system and resulted into elevated Hsp70 level and lipid peroxidation of tissues. These information will help to develop suitable culture strategy for this economically important food fish. The present study also confirms the suitability of the studied parameters as efficient thermal biomarkers for fish.

Acknowledgments Authors are thankful to Prof. P. Mittal, Satyawati College, University of Delhi, for his significant contribution in the statistical analysis and interpretation.

Authors' contributions RC and JGS designed the study. TA, SPS, JGS, and RC cultured the fish and analyzed samples. RC and JGS prepared and revised the manuscript. RC, SPS, and TA prepared the tables and graphs. RC performed statistical analysis.

Funding This study has been carried out with the financial support from Indian Council of Agricultural Research, ICAR, New Delhi (NFBSFARA project, AS-2001/2010-11).Data availabilityAll data related to this study are included in the manuscript.

Compliance with ethical standards The guidelines of Animal Ethics Committee (IAEC), Department of Zoology, University of Delhi, Delhi, India, was followed for the study.

Conflict of interest The authors declare that they have no conflict of interest.

References

Ahmad T, Singh SP, Khangembam BK, Sharma JG, Chakrabarti (2014) Food consumption and digestive enzyme activity of *Clarias batrachus* exposed to various temperatures. Aquac Nutr 20:265–272

Almroth BC, Asker N, Wassmur B, Rosengren A, Jutfelt F, Gräns A et al (2015) Warmer water temperature results in oxidative damage in an Antarctic fish, the bald notothen. J Exp Mar Biol Ecol 468:130–137

Anderson DP, Siwicki AK (1995) Basic haematology and serology for fish health programs. In: Shariff M, Arthur JR, Subasinghe RP (eds) Diseases in Asian aquaculture II, Manila. Fish Health Section, Asian Fisheries Society, Manila, Philippines, pp 185–202

Andrew PJ, Mayer B (1999) Enzymatic function of nitric oxide synthases. Cardiovasc Res 43:521–531

Bagnyukova TV, Chahrak OI, Lushchak VI (2006) Coordinated response of goldfish antioxidant defenses to environmental stress. Aquat Toxicol 78:325–331

Bagnyukova TV, Lushchak OV, Storey KB, Lushchak VI (2007) Oxidative stress and antioxidant defense responses by goldfish tissues to acute change of temperature from 3 to 23 °C. J Therm Biol 32:227–334

Basu N, Todgham AE, Ackerman PA, Bibeau MR, Nakano K, Schulte PM, Iwama GK (2002) Heat shock protein genes and their functional significance in fish. Gene 295:173–183

Beevi SS, Narasu ML, Gowda BB (2010) Polyphenolics profile, antioxidant and radical scavenging activity of leaves and stem of *Raphanus sativus* L. Plant Food Hum Nutr 65:8–17

Bendiksen EÅ, Jobling M, Arnesen AM (2002) Feed intake of Atlantic salmon parr *Salmo salar* L. in relation to temperature and feed composition. Aquac Res 33:525–532

Billar JD, Takashi LS (2018) Oxidative stress and fish immune system: phagocytosis and leukocyte respiratory burst activity. An Acad Bras Cienc 90:3403–3414

Blokhina O, Virolainen E, Fagerstedt KV (2003) Antioxidants, oxidative damage and oxygen deprivation stress: a review. Ann Bot 91:179–194

Bocchetti R, Lamberti CV, Pisanelli B, Razzetti EM, Maggi C, Catalano B, Sesta G, Martuccio G, Gabellini M, Regoli F (2008) Seasonal variations of exposure biomarkers, oxidative stress responses and cell damage in the clams, *Tapes philippinarum*, and mussels, *Mytilus galloprovincialis*, from Adriatic Sea. Mar Environ Res 66:24–26

Brander KM (1994) Patterns of distribution, spawning, and growth in North Atlantic cod: the utility of inter-regional comparisons. ICES Mar Sci Symp 198:406–413

Brown JA, Pepin P, Methven DA, Somerton DC (1989) The feeding, growth and behaviour of juvenile cod, *Gadus morhua* L, in cold environments. J Fish Biol 35:373–380

Cardoso CM, Sartorio PV, Machado AS, Vignardi CP, Rojas DC, Passos MJ, Rocha AJ, Van Ngan P, Gomes V (2015) Hsp70 and p53 expressions and behavior of juvenile pompano, *Trachinotus carolinus* (Perciformes, Carangidae), at controlled temperature increase. J Exp Mar Biol Ecol 470:34–42

Chakrabarti R, Sharma JG (2005) Digestive physiology of fish larvae during ontogenic development: a brief overview. Indian J Anim Sci 75:1337–1347

Claiborne A (1985) Catalase activity. In: Greenwald RA (ed) CRC Handbook of methods in oxygen radical research. CRC Press Inc, Boca Raton, pp 283–284

Clark DS, Brown JA, Goddard SJ, Moir J (1995) Activity and feeding behaviour of Atlantic cod (*Gadus morhua*) in sea pens. Aquaculture 131:49–57

Coles BF, Kadlubar FF (2003) Detoxification of electrophilic compounds by glutathione S-transferase catalysis: determinants of individual response to chemical carcinogens and chemotherapeutic drugs? Biofactors 17:115–130

Dalvi RS, Pal AK, Tiwari LR, Das T, Baruah K (2009) Thermal tolerance and oxygen consumption rates of the catfish *Horabagrus brachysoma* (Günther) acclimated to different temperatures. Aquaculture 295:116–119

Davies MJ (2011) Myeloperoxidase-derived oxidation: mechanisms of biological damage and its prevention. J Clin Biochem Nutr 48:8–19

Deane EE, Woo N (2005) Expression studies on glucose-6-phosphate dehydrogenase in sea bream: effects of growth hormone, somatostatin, salinity and temperature. J Exp Zool A 303:676–688

El-Ansary AK, Bacha AB, Kotb M (2012) Etiology of autistic features: the persisting neurotoxic effects of propionic acid. J Neuroinflammation 9:74 http://www.jneuroinflammation.com/content/9/1/74

Froese R, Pauly D (2019) Fish base. World Wide Web electronic publication. Available at: http://www.fishbase.org (12/2019) (accessed on July 20, 2020)

Fu D, Chen J, Zhang Y, Yu Z (2011) Cloning and expression of a heat shock protein (HSP) 90 gene in the haemocytes of *Crassostrea hongkongensis* under osmotic stress and bacterial challenge. Fish Shellfish Immunol 31:118–125

Gordon CJ (2005) Temperature and toxicology: an integrative, comparative, and environmental approach. CRC Press, Boca Raton

Grim JM, Miles DRB, Crockett EL (2010) Temperature acclimation alters oxidative capacities and composition of membrane lipids without influencing activities of enzymatic antioxidants or susceptibility to lipid peroxidation in fish muscle. J Exp Biol 213:445–452

Habig WH, Pabst MJ, Jakoby WB (1974) Glutathione S-transferases the first enzymatic step in mercapturic acid formation. J Biol Chem 249:7130–7139

Haniffa MA, Marimuthu K, Nagarajan M, Arokiaraj AJ, Kumar D (2004) Breeding behaviour and parental care of the induced breed spotted murrel *Channa punctata* under captivity. Curr Sci 86:1375–1376

Haugland GT, Jordal AEO, Wergeland HI (2012) Characterization of small, mononuclear blood cells from salmon having high phagocytic capacity and ability to differentiate into dendritic like cells. PLoS One 7:e49260

Haynes W (2013) Tukey's Test. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H (eds) Encyclopedia of systems biology. Springer, New York. https://doi.org/10.1007/978-1-4419-9863-7_1212

Hossen MA, Paul AK, Hossain MDY, Ohtomi J, Sabbir W, Rahman O, Jasmin J, Khan MDN, Islam MDAR, Khatun D, Kamruzzaman SK (2017) Estimation of biometric indices for snakehead *Channa punctata* (Bloch, 1793) through multi-model inferences. JJBS 12:197–202

Iwama GK, Vijayan MM, Forsyth RB, Ackerman PA (1999) Heat shock proteins and physiological stress in fish. Am Zool 39:901–909

Iwama GK, Afonso LO, Todgham A, Ackerman P, Nakano K (2004) Are hsps suitable for indicating stressed states in fish? J Exp Biol 207:15–19

Jollow DJ, Mitchell JR, Zampaglione NA, Gillette JR (1974) Bromobenzene-induced liver necrosis. Protective role of glutathione and evidence for 3, 4-bromobenzene oxide as the hepatotoxic metabolite. Pharmacology 11:151–169

Kaur M, Atif F, Ali M, Rehman H, Raisuddin S (2005) Heat stress-induced alterations of antioxidants in the freshwater fish *Channa punctata* Bloch. J Fish Biol 67:1653–1665

Killen SS (2014) Growth trajectory influences temperature preference in fish through an effect on metabolic rate. J Anim Ecol 83:1513–1522

Kinkade JM Jr, Pember SO, Barnes KC, Shapira R, Spitznagel JK, Martin LE (1983) Differential distribution of distinct forms of myeloperoxidase in different azurophilic granule subpopulations from human neutrophils. Biochem Biophys Res Commun 114:296–303

Klebanoff SJ (1999) Myeloperoxidase. Proc Assoc Am Physicians 111:383–389

Klebanoff SJ (2005) Myeloperoxidase: friend and foe. J Leukoc Biol 77:598–625

Kohen R, Nyska A (2002) Invited review: oxidation of biological systems: oxidative stress phenomena, antioxidants, redox reactions, and methods for their quantification. Toxicol Pathol 30:620–650

Lee DU, Kang YJ, Park MK, Lee YS, Seo HG, Kim TS, Kim CH, Chang KC (2003) Effects of 13-alkyl-substituted berberine alkaloids on the expression of *COX-II, TNF-α, iNOS*, and *IL-12* production in LPS-stimulated macrophages. Life Sci 73:1401–1412

Liu Y, Ma D, Zhao C, Xiao Z, Xu S, Xiao Y, Wang Y, Liu Q, Li J (2017) The expression pattern of hsp70 plays a critical role in thermal tolerance of marine demersal fish: multilevel responses of *Paralichthys olivaceus* and its hybrids (*P. olivaceus*♀ × *P. dentatus*♂) to chronic and acute heat stress. Mar Environ Res 129:386–395

Liu Z-M, Zhu X-L, Lu J, Cai W-J, Ye Y-P, Yao-Ping Lv Y-P (2018) Effect of high temperature stress on heat shock protein expression and antioxidant enzyme activity of two morphs of the mud crab *Scylla paramamosain*. Comp Biochem Physiol A 223:10–17

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. J Biol Chem 193:265–275

Lushchak VI (2011) Environmentally induced oxidative stress in aquatic animals. Aquat Toxicol 101:13–30

Lushchak VI, Bagnyukova TV (2006a) Temperature increase results in oxidative stress in goldfish tissues. 1. Indices of oxidative stress. Comp Biochem Physiol C Toxicol Pharmacol 143:30–35

Lushchak VI, Bagnyukova TV (2006b) Temperature increase results in oxidative stress in goldfish tissues. 2. Antioxidant and associated enzymes. Comp Biochem Physiol C Toxicol Pharmacol 143:36–41

Mariana S, Alfons, Gamal B (2019) Impact of heat stress on the immune response of fishes. J Sur Fish Sci 5:149–159

Marimuthu K, Arokiaraj J, Haniffa MA (2009) Effect of diet quality on seed production of the spotted snakehead *Channa punctatus* (Bloch). Am-Eurasian J Sustain Agric 3: 344–347

Mohandas J, Marshall JJ, Duggin GG, Horvath JS, Tiller DJ (1984) Differential distribution of glutathione and glutathione-related enzymes in rabbit kidney: possible implications in analgesic nephropathy. Biochem Pharmacol 33: 1801–1807

Nie H, Liu L, Huo Z, Chen P, Ding J, Yang F, Yan X (2017) The HSP70 gene expression responses to thermal and salinity stress in wild and cultivated Manila clam *Ruditapes philippinarum*. Aquaculture 470:149–156

Noori S (2012) An overview of oxidative stress and antioxidant defensive system. Sci Rep 1:1–9

Ohkawa H, Ohishi N, Yagi K (1979) Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal Biochem 95:351–358

Palić D, Herolt DM, Andreasen CB (2006) Anesthetic efficacy of tricaine methanesulfonate, metomidate and eugenol: effects on plasma cortisol concentration and neutrophil function in fathead minnows (*Pimephales promelas* Rafinesque, 1820). Aquaculture 254:675–685

Quade MJ, Roth JA (1997) A rapid, direct assay to measure degranulation of bovine neutrophil primary granules. Vet Immunol Immunopathol 58:239–248

Rabergh CMI, Airaksinen S, Soitamo A, Bjorklund HV, Johansson T, Nikinmaa M, Sistonen L (2000) Tissue-specific expression of zebrafish (*Danio rerio*) heat shock factor 1 mRNAS in response to heat stress. J Exp Biol 203: 1817–1824

Sales CF, Lemos FS, Morais RD, Thomé RG, Santos HB, Pinheiro AP, Bazzoli N, Rizzo E (2019) Thermal stress induces heat shock protein 70 and apoptosis during embryo development in a Neotropical freshwater fish. Reprod Fertil Dev 31:547–556. https://doi.org/10.1071/RD18217

Saranyan PV, Ross NW, Benfey TJ (2017) Erythrocyte heat shock protein responses to chronic (*in vivo*) and acute (*in vitro*) temperature challenge in diploid and triploid salmonids. Comp Biochem Physiol A 206:95–104

Schairer DO, Chouake J, Nosanchuk JD, Friedman JM (2012) The potential of nitric oxide releasing therapies as antimicrobial agents. Virulence 3:271–279

Scott MA, Locke M, Buck LT (2003) Tissue-specific expression of inducible and constitutive Hsp70 isoforms in the western painted turtle. J Exp Biol 206:303–311

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality. Biometrika 52:591–611

Sharma JG, Singh SP, Mittal P, Chakrabarti R (2014) Impact of temperature gradient on the Indian major carp *Catla catla* larvae. Proc Natl Acad Sci India Sect B Biol Sci 86:269–273

Sharma JG, Singh SP, Chakrabarti R (2017) Effect of temperature on digestive physiology, immune-modulatory parameters, and expression level of Hsp and LDH genes in *Catla catla* (Hamilton, 1822). Aquaculture 479:134–141

Shin MK, Park HR, Yeo WJ, Han KN (2018) Effects of thermal stress on the mRNA expression of SOD, HSP90, and HSP70 in the spotted sea bass (*Lateolabrax maculatus*). Ocean Sci J 53:43–52

Singh SP, Sharma JG, Ahmad T, Chakrabarti R (2013) Effect of water temperature on the physiological responses of Asian catfish *Clarias batrachus* (Linnaeus 1758). Asian Fish Sci 26:26–38

Singh KM, Sharma JG, Chakrabarti R (2015) Simulation study of natural UV-B radiation on *Catla catla* and its impact on physiology, oxidative stress, Hsp70 and DNA fragmentation. Photochem Photobiol 149:156–163

Somero GN (2002) Thermal physiology and vertical zonation of intertidal animals: optima, limits, and costs of living. Integr Comp Biol 42:780–789

Talwar PK, Jhingran AG (1991) Inland fishes of India and adjacent countries. vol 1. A.A. Balkema, Rotterdam, p 541

Verlecar XN, Jena KB, Chainy GB (2007) Biochemical markers of oxidative stress in *Perna viridis* exposed to mercury and temperature. Chem Biol Interact 167:219–226

Vinagre C, Madeira D, Narciso L, Cabral HN, Diniz M (2012) Effect of temperature on oxidative stress in fish: lipid peroxidation and catalase activity in the muscle of juvenile seabass, *Dicentrarchus labrax*. Ecol Indic 23:274–279

Wang J, Wei Y, Li X, Cao H, Xu M, Dai J (2007) The identification of heat shock protein genes in goldfish (*Carassius auratus*) and their expression in a complex environment in Gaobeidian Lake, Beijing, China. Comp Biochem Physiol C Toxicol Pharmacol 145:350–362

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

End to End Binarized Neural Networks for Text Classification

Kumar Shridhar[1]*, **Harshil Jain**[2]*, **Akshat Agarwal**[3]*, **Denis Kleyko**[4,5]
[1]NeuralSpace, London
[2]Computer Science and Engineering, IIT Gandhinagar, Gujarat, India
[3]Electrical Engineering, Delhi Technological University, Delhi, India
[4]Redwood Center for Theoretical Neuroscience, University of California, Berkeley
[5]Intelligent Systems Lab, Research Institutes of Sweden
kumar@neuralspace.ai, jain.harshil@iitgn.ac.in,
akshat.agarwal0311@gmail.com, denis.kleyko@ri.se

Abstract

Deep neural networks have demonstrated their superior performance in almost every Natural Language Processing task, however, their increasing complexity raises concerns. A particular concern is that these networks pose high requirements for computing hardware and training budgets. The state-of-the-art transformer models are a vivid example. Simplifying the computations performed by a network is one way of addressing the issue of the increasing complexity. In this paper, we propose an end to end binarized neural network for the task of intent and text classification. In order to fully utilize the potential of end to end binarization, both the input representations (vector embeddings of tokens statistics) and the classifier are binarized. We demonstrate the efficiency of such a network on the intent classification of short texts over three datasets and text classification with a larger dataset. On the considered datasets, the proposed network achieves comparable to the state-of-the-art results while utilizing $\sim$ 20-40% lesser memory and training time compared to the benchmarks.

1 Introduction

In recent years, deep neural networks have achieved great success in a variety of domains, but the networks are becoming more and more computationally expensive due to their ever-growing size. This tendency has been noticed in (Strubell et al., 2019; Schwartz et al., 2019) and it has been recommended that academia and industry researchers should draw their attention towards more computationally efficient methods. At the same time, many important application areas such as chatbots, IoT devices, mobile devices, and other types of power-constrained and resource-constrained platforms require solutions that would be highly computationally and memory efficient. Such use-cases limit the potential use of the state-of-the-art deep networks. One viable solution is the transformation of these high-performance neural networks to a more computationally efficient architecture. Recently, Binarized Convolutional Neural Networks (BNN) (Hubara et al., 2016) have been developed where both weights and activations are restricted to $\{+1, -1\}$. BNN is a highly computationally efficient network with a much lower memory footprint. Tasks like language modeling (Zheng and Tang, 2016) were performed using binarized neural networks, but, to the best of our knowledge, in the area of text classification, no end to end trainable binarized architectures have been demonstrated yet.

In this paper, we introduce an architecture for the tasks of intent and text classifications that fully utilizes the power of binary representations. The input representations are tokenized and embedded in binary high-dimensional (HD) vectors forming distributed representations using the paradigm known as hyperdimensional computing (Kanerva, 2009). The binary input representations are used for training an end to end BNN classifier for intent classification. Classification performance-wise, the binarized architecture achieves results comparable to the state-of-the-art on several standard intent classification datasets. The efficiency of the proposed architecture is shown in terms of its time and memory complexity relative to non-binarized architectures.

2 Proposed Method

Figure 1 presents a schematic overview of the architecture. Given an input text document D, we first pre-process the document. The pre-processed document is then tokenized into the corresponding tokens $< T_1, T_2, ..., T_n >$, which are used as an

* The authors contributed equally to this research and work was done at NeuralSpace

Proceedings of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing, pages 29–34
Online, November 20, 2020. ©2020 Association for Computational Linguistics

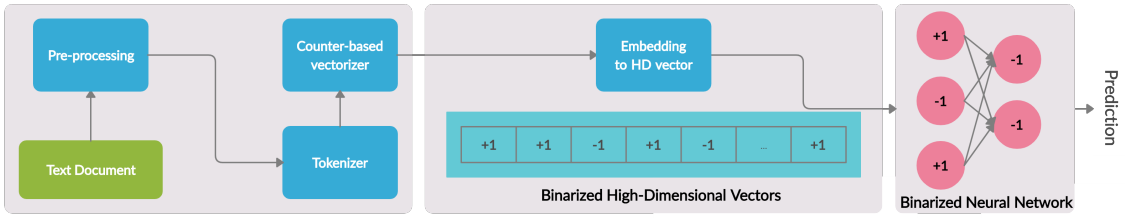


Figure 1: A schematic diagram of the end to end binarized classification architecture for text classification.

input to a count-based vectorizer. The representation of vectorizers, which is sparse and localist, is embedded into an HD vector (distributed representation) using hyperdimensional computing. HD vector representing the counter's content can be binarized. It is used as an input to a classifier. The primary classifier studied in this work is BNN, but other classifiers are also considered for benchmarking.

2.1 High-Dimensional embedding of vectorized representations

In order to reduce the dimensionality of representations, we use hyperdimensional computing (Kanerva, 2009). First, each unique token T_i is assigned with a random d-dimensional bipolar HD vector, where d would be a hyperparameter of the method. HD vectors are stored in the item memory, which is a matrix $\mathbf{H} \in [d \times n]$, where n is the number of tokens. Thus, for a token T_i there is an HD vector $\mathbf{H}_{T_i} \in \{-1, +1\}^{[d \times 1]}$. To construct composite representations from the atomic HD vectors stored in $\mathbf{H}$, hyperdimensional computing defines three key operations: permutation (ρ), binding ($\odot$, implemented via element-wise multiplication), and bundling ($+$, implemented via element-wise addition) (Kanerva, 2009). The bundling operation allows storing information in HD vectors (Frady et al., 2018). The three operations above allow embedding vectorized representations based on n-gram statistics into an HD vector (Joshi et al., 2016).

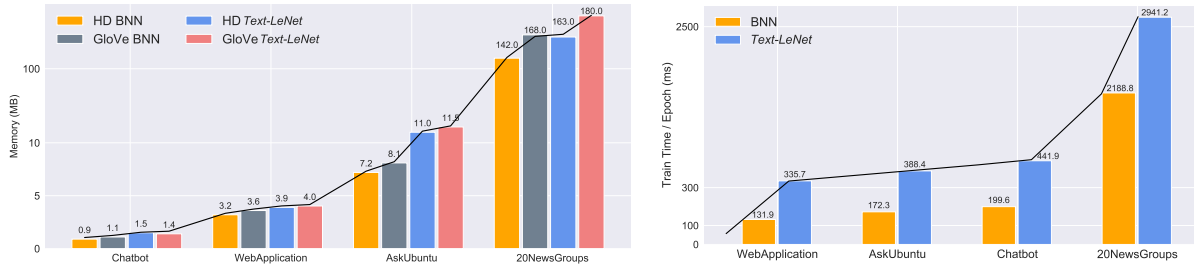
We first generate $\mathbf{H}$, which has an HD vector for each token. The permutation operation ρ is applied to $\mathbf{H}_{T_j}$ j times ($\rho^j(\mathbf{H}_{T_j})$) to represent a relative position of token T_j in an n-gram. A single HD vector corresponding to an n-gram (denoted as $\mathbf{m}$) is formed using the consecutive binding of permuted HD vectors $\rho^j(\mathbf{H}_{T_j})$ representing tokens in each position j of the n-gram. For example, the trigram '#he' will be embedded to an HD vector as follows: $\rho^1(\mathbf{H}_\#) \odot \rho^2(\mathbf{H}_\mathrm{h}) \odot \rho^3(\mathbf{H}_\mathrm{e})$. In general, the process of forming HD vector of an n-gram is $\mathbf{m} = \prod_{j=1}^{n} \rho^j(H_{T_j})$, where T_j is token in jth position of the n-gram; the consecutive binding operations applied to n HD vectors are denoted by $\prod$. Once it is known how to form an HD vector for an individual n-gram, embedding the n-gram statistics into an HD vector $\mathbf{h}$ is achieved by bundling together all n-grams observed in the document:

$$\mathbf{h} = [\sum_{i=1}^{k} f_i \mathbf{m}_i = \sum_{i=1}^{k} f_i \prod_{j=1}^{n} \rho^j(H_{T_j})],$$

where k is the total number of unique n-grams; f_i is the frequency of ith n-gram and $\mathbf{m}_i$ is the HD vector of ith n-gram; $\sum$ denotes the bundling operation when applied to several HD vectors; $[*]$ denotes the binarization operation, which is implemented via the sign function. The usage of $[*]$ is optional, so we can either obtain binarized or non-binarized $\mathbf{h}$. If $\mathbf{h}$ is non-binarized, its components will be integers in the range $[-k, k]$, but these extreme values are highly unlikely since HD vectors for different n-grams are quasi-orthogonal, which means that in the simplest (but not practical) case when all n-grams have the same probability the expected value of a component in $\mathbf{h}$ is 0. Due to the use of $\sum$ for representing n-gram statistics, two HD vectors embedding two different n-gram statistics might have very different amplitudes if the frequencies in these statistics are very different. When HD vectors $\mathbf{h}$ are binarized, this issue is addressed. In the case of non-binarized HD vectors, we address it by using the cosine similarity, which is imposed by normalizing each $\mathbf{h}$ by its ℓ_2 norm; thus, all $\mathbf{h}$ have the same norm, and their dot product is equivalent to their cosine similarity.

2.2 Binarized Neural Networks

Based on the work of (Hubara et al., 2016), we construct BNNs capable of working with representations of texts. To take the full advantage of binarized HD vectors, we constraint the weights and



(a) Memory Comparison

(b) Train Time per epoch Comparison

Figure 2: (a) shows the memory comparisons for all 4 datasets using HD *Text-LeNet*, HD BNN, GloVe *Text-LeNet*, GloVe BNN and (b) shows the training time per epoch comparison for all 4 datasets using BNN and *Text-LeNet*

	Chatbot		AskUbuntu		WebApplication		20NewsGroups	
Tokenizers	*Text-LeNet*	BNN	*Text-LeNet*	BNN	*Text-LeNet*	BNN	*Text-LeNet*	BNN
Word	**0.80**	0.73	0.51	**0.79**	0.56	**0.78**	0.54	**0.56**
SemHash	**0.94**	0.90	**0.87**	0.84	0.79	**0.83**	**0.78**	0.69
BPE	**0.80**	0.58	0.54	**0.67**	0.52	**0.75**	0.38	**0.42**
Char BPE	**0.92**	0.81	**0.76**	**0.76**	**0.55**	0.53	**0.55**	0.48
SentencePiece	0.80	**0.99**	0.70	**0.72**	0.50	**0.70**	0.41	**0.43**
BERT	**0.89**	0.88	**0.72**	0.71	0.70	**0.77**	**0.60**	**0.60**

Table 1: F_1 performance comparison of binarized *Text-LeNet* (BNN) architecture with non-binarized *Text-LeNet* for the task of intent classification on various datasets.

activations of the network layers to be $\{+1, -1\}$. This constraint is highly efficient in terms of hardware and memory, as bit-wise operations are used instead of multiply-accumulate operations. For example, a multiplication on binary values can be performed using an XNOR logical operation.

The vectorized representations of tokens embedded into HD vectors are binarized with all values $\{+1, -1\}$. In the case of HD vectors, we binarize the result of the bundling operation using the sign function.

Similarly, the sign function is used in the BNN for every weight or activation to restrict them into $\{+1, -1\}$ as follows:

$$b(x) = [x] = \text{sign}(x) = \begin{cases} +1 & \text{if } x \geq 0, \\ -1 & \text{otherwise} \end{cases} \quad (1)$$

where, x can be any weight or activation value.

We further define a convolutional 1D layer that creates a convolution kernel that is convolved with the input HD vector over a single spatial dimension to produce a tensor of outputs. Since gradient descent methods make small changes to the value of the weights, which cannot be done with binary values, we use the straight-through estimator idea, as mentioned in (Yin et al., 2019). We also define a value over which we clip the gradients in the backward pass:

$$\frac{\delta b(x)}{\delta x} = \begin{cases} +1 & \text{if } |x| < \text{clip value}, \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

This ensures that the entire architecture is end to end trainable using gradient descent optimization.

3 Empirical Analysis

3.1 Datasets

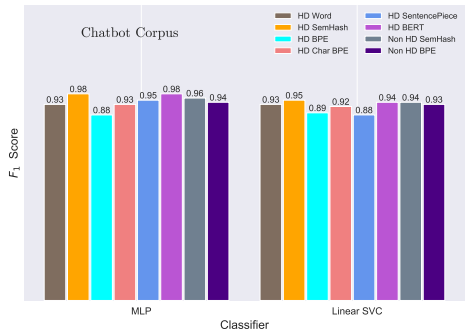
All the experiments are performed on four datasets, namely: the *Chatbot Corpus* (Chatbot), the *Ask Ubuntu Corpus* (AskUbuntu), the *Web Applications Corpus* (WebApplication), and the *20 News Groups Corpus* (20NewsGroups) (Braun et al., 2017).

3.2 Results and Discussions

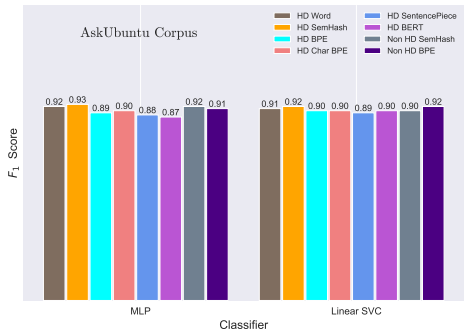
For CNN-based architecture, 5 hidden layers were used: 3 convolutional 1D layers followed by 2 dense layers. Due to its resemblance to the original LeNet architecture (LeCun et al., 1998), we refer to this architecture as *Text-LeNet*. We compare the results of binarized HD vectors with the binarized *Text-LeNet* (BNN) architecture as the classifier against non-binarized HD vectors with non-binarized *Text-LeNet*. The F_1 scores are compared in Table 1 where BNN performed equally well to a *Text-LeNet* architecture while being 20%

Datasets	Binarized GloVe	Binarized SemHash	Binarized HD vectors
Chatbot	0.74	0.91	**0.99**
AskUbuntu	0.86	**0.87**	0.84
WebApplication	0.66	0.80	**0.83**
20NewsGroups	0.62	0.64	**0.69**

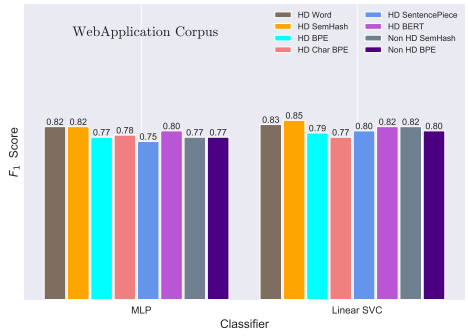
Table 2: F_1 performance comparison of Binarized GloVe vectors, Binarized SemHash vectors and Binarized HD vectors. All vectorizers use the same binarized *Text-LeNet* architecture as classifier.



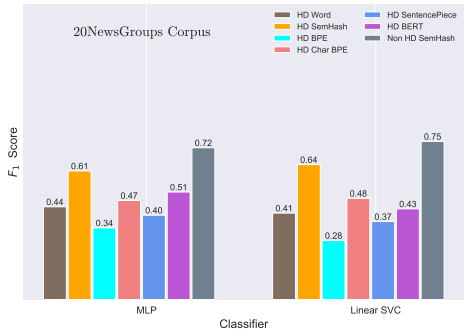
(a) Chatbot Corpus



(b) AskUbuntu Corpus



(c) WebApplication Corpus



(d) 20NewsGroups Corpus

Figure 3: (a), (b), (c) and (d) show the F_1 score comparison of MLP and Linear SVC classifier with HD and non-HD based tokenizers on Chatbot, AskUbuntu, WebApplication and 20NewsGroups corpus respectively.

to 40% more memory efficient, as shown in Figure 2 (a). Note that due to the specifics of implementation, BNNs use 32 bit float values as *Text-LeNet*. The memory efficiency of BNNs can be further improved by 4x when 8-bit representations are used and up to 32x if a single bit representations are used. However, the hardware limitations prevented us from going to that extreme. On the performance side, BNNs outperforms the *Text-LeNet* for AskUbuntu and WebApplication datasets on 4 out of 6 tokenizers. The results reported in Table 1 used 512 dimensional HD vectors for Chatbot, AskUbuntu, and WebApplication corpus, while 1,024 dimensional HD vectors were used for the 20NewsGroups dataset.

One thing to note here is that *Text-LeNet* also used HD vectors with the mentioned tokenizers, but the HD vectors were non-binarized. HD vectors in itself are already faster and much more efficient than counter-based representations, as shown in (Alonso et al., 2020). When experimenting with other embedding methods like GloVe, the training was significantly slower; therefore, HD vectors were used for all the experiments. In addition to that, using the binarized classifier (BNN) further improved the training time up to 50% per epoch when compared to non-binarized classifier on all four datasets, as shown in Figure 2 (b). Furthermore, when compared to GloVe embeddings with *Text-LeNet*, HD BNN used around 20 - 40% lesser

Platform	Chatbot	AskUbuntu	WebApplication	Average
Botfuel	0.98	0.90	0.80	0.89
Luis	0.98	0.90	0.81	0.90
Dialogflow	0.93	0.85	0.80	0.86
Watson	0.97	0.92	0.83	0.91
Rasa	0.98	0.86	0.74	0.86
Snips	0.96	0.83	0.78	0.86
Recast	**0.99**	0.86	0.75	0.87
TildeCNN	**0.99**	0.92	0.81	0.91
FastText	0.97	0.91	0.76	0.88
SemHash (Shridhar et al., 2019)	0.96	0.92	**0.87**	**0.92**
BPE	0.95	**0.93**	0.85	0.91
HD vectors (Alonso et al., 2020)	0.97	0.92	0.82	0.90
Binarized HD vectors with the best classifier	0.98	**0.93**	0.84	**0.92**
HD *Text-LeNet*	0.94	0.87	0.79	0.88
HD BNN	**0.99**	0.84	0.83	0.91

Table 3: F_1 score comparison of various platforms on intent classification datasets of short texts with methods used in the paper. Some results are taken from (Alonso et al., 2020)

memory for all the intent classification datasets.

We also benchmarked the binarized HD vectors with binarized 300-dimensional GloVe vectors and the binarized version of counter-based representation for SemHash tokenizer (Alonso et al., 2020) for all the datasets. Table 2 summarizes the results of the comparison. All the binarized representations were trained with the same BNN classifier. Binarized HD vectors performed significantly better than other binarized methods outperforming binarized GloVe by 4 - 25% and binarized SemHash by 2 - 8% on 2 out of 3 smaller intent classification datasets and achieved comparable results for AskUbuntu dataset. The trend continued for 20NewsGroups with binarized HD achieving 5 - 7% better F_1 scores. Note that for the SemHash counter-based vectorizer, we put a sign function $\text{sign}(x) = +1$ for $x > 0$ and -1 otherwise.

In Figure 3, MLP and Linear SVC with all the tokenizers with HD vectors as representation are compared with MLP and Linear SVC classifiers with SemHash tokenizers and counter-based vectorizer as representation from (Alonso et al., 2020). The F_1 score is comparable to the state-of-the-art for both MLP and SVC. For all small intent classification datasets, binarized HD vectors have achieved better results than non-HD vectors. The proposed architecture beats the non-HD baselines by $+2\%$ for AskUbuntu and Chatbot Corpus, and $+5\%$ for WebApplication Corpus. However, for 20NewsGroups, the results of binarized HD Vectors are lower than non-HD Vectors. This is mainly due to the large size of the dataset, and simple classifiers like LinearSVC failed to perform with just binarized values. The results for all the other classifiers are provided in the Appendix.

Table 3 compares the F_1 scores of various platforms on the intent classification datasets. We report the results of binarized HD vectors with the best classifiers from one of the nine classifiers mentioned (Binarized HD vectors with the best classifier), non-binarized HD vectors with *Text-LeNet* (HD *Text-LeNet*) and binarized HD vectors with binarized *Text-LeNet* (HD BNN). Our end to end binarized architecture (HD BNN) achieved the state-of-the-art results for the Chatbot dataset. The approach where only HD vectors were binarized (binarized HD vectors with the best classifier) achieved the state-of-the-art results for the AskUbuntu dataset. The results on the WebApplication dataset are comparable to the state-of-the-art (0.87 with SemHash): 0.84 for binarized HD vectors with the best classifier and 0.83 for HD BNN. The average performance of both binarized HD vectors with the best classifier (0.92) and HD BNN (0.91) was also comparable to the best non-binarized approach (0.92).

4 Conclusion

In this work, we show that it is possible to achieve comparable to the state-of-the-art results while using the binarized representations of all the components of the text classification architecture. This allows exploring the effectiveness of binary representations both for reducing the memory footprint of the architecture and for increasing the energy-efficiency of the inference phase due to the effectiveness of binary operations. This work takes a step towards enabling NLP functionality on resource-constrained devices.

References

P. Alonso, K. Shridhar, D. Kleyko, E. Osipov, and M. Liwicki. 2020. HyperEmbed: Tradeoffs between Resources and Performance in NLP Tasks with Hyperdimensional Computing Enabled Embedding of n-gram Statistics. *arXiv:2003.01821*.

D. Braun, A. Hernandez-Mendez, F. Matthes, and M. Langen. 2017. Evaluating Natural Language Understanding Services for Conversational Question Answering Systems. In *Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 174–185.

E. P. Frady, D. Kleyko, and F. T. Sommer. 2018. A Theory of Sequence Indexing and Working Memory in Recurrent Neural Networks. *Neural Computation*, 30:1449–1513.

L. Geiger and P. Team. 2020. Larq: An Open-Source Library for Training Binarized Neural Networks. *Journal of Open Source Software*, 5(45):1746.

G. Hinton, N. Srivastava, and K. Swersky. 2012. Neural Networks for Machine Learning Lecture 6a Overview of Mini-batch Gradient Descent.

I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio. 2016. Binarized Neural Networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1–9.

S. Ioffe and C. Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167.

A. Joshi, J. T. Halseth, and P. Kanerva. 2016. Language Geometry Using Random Indexing. In *Quantum Interaction (QI)*, pages 265–274.

P. Kanerva. 2009. Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognitive Computation*, 1(2):139–159.

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

R. Schwartz, J. Dodge, N. Smith, and O. Etzioni. 2019. Green ai. *arXiv preprint arXiv:1907.10597*.

K. Shridhar, A. Dash, A. Sahu, G. Grund Pihlgren, P. Alonso, V. Pondenkandath, G. Kovacs, F. Simistira, and M. Liwicki. 2019. Subword Semantic Hashing for Intent Classification on Small Datasets. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–6.

E. Strubell, A. Ganesh, and A. McCallum. 2019. Energy and Policy Considerations for Deep Learning in NLP. In *57th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 3645–3650.

P. Yin, J. Lyu, S. Zhang, S. Osher, Y. Qi, and J. Xin. 2019. Understanding Straight-through Estimator in Training Activation Quantized Neural Nets. *arXiv:1903.05662*.

W. Zheng and Y. Tang. 2016. Binarized Neural Networks for Language Modeling. *Technical Report cs224d, Stanford University*.

Journal of Polymer Research (2020) 27:358
https://doi.org/10.1007/s10965-020-02323-x

ORIGINAL PAPER

Enhanced piezoelectricity properties of reduced graphene oxide (RGO) loaded polyvinylidene fluoride (PVDF) nanocomposite films for nanogenerator application

Abhishek Anand[1] · Deshraj Meena[1,2] · Kajal Kumar Dey[3] · Mukesh Chander Bhatnagar[1]

Received: 17 June 2019 / Accepted: 8 October 2020
© The Polymer Society, Taipei 2020

Abstract
Flexible and lightweight films of Polyvinylidene fluoride (PVDF) loaded with nanosheets of reduced graphene oxide (RGO) have been prepared via solution casting method. The structure of RGO nanosheets have been examined through X-ray diffraction (XRD), Transmission electron microscopy (TEM) and Raman spectroscopy technique. FTIR studies revealed that the enhancement of β-phase from 53 to 70% has been observed by adding RGO nanosheets in PVDF matrix. The frequency dependent dielectric constant and dielectric loss have been studied. Ultraviolet–visible spectra show a decrease in optical band gap with increases RGO concentration in nanocomposite films. The ferroelectric behavior of RGO loaded PVDF nanocomposite films showed remarkable improvement in remnant polarization values as compared to bare PVDF. The piezo-response of RGO loaded nanocomposite films showed significant enhancement in output voltage from 0.886 to 1.915 V. On exposure to ultraviolet–visible light, RGO loaded nanocomposite PVDF films showed about 13.8% enhancement in piezo- voltage with action of same pressure.

Keywords PVDF · Piezoelectric · β-phase · RGO · Nanocomposite

Introduction

The piezoelectric energy conversion from mechanical energy to electrical energy referred as the 'Direct piezoelectric effect' is one of the most versatile phenomenon in an energy harvesting process [1]. Therefore, Energy harvesting from irregular air flow/vibration, ultrasonic waves, and body movement has created enormous interest in the scientific community because of the easily available and environmental factors which exist enormously in our daily activity. In the past few decades, a piezoelectric material has an utmost demand in various devices and application in modern society. With their intrinsic electromechanical properties, applied strain produces an electrical signal, whereas electrical input induces mechanical deformation. Piezoelectric materials devices are components of choice in strain sensors, self-powered wireless sensors, acoustic and ultrasonic devices, precision positioning in scanning probe microscopes, fuel injection systems, active damping control, and adaptive control systems. For that designing of flexible piezoelectric nanogenerator (PNG), using electroactive polar Poly(vinyldene fluoride) (PVDF) polymer is a very attractive polymer which having dielectric, piezoelectric, ferroelectric, pyroelectric and other excellent physical properties [2].

Poly(vinyldene fluoride) PVDF $(-CH_2-CF_2-)_n$ is a semi-crystalline polymer and possesses five different crystalline phases α, β, γ, δ and ε in which α phase is non-polar whereas β, γ, δ and ε are polar phases. Although PVDF has higher piezoelectric coefficient, d_{33} than other polymers yet it is not preferred suitable for the application of energy harvesting because of low piezoresponse [3]. The non-polar α phase,

✉ Abhishek Anand
abhishek.allen.anand@gmail.com

Deshraj Meena
meenadesh707@gmail.com

Mukesh Chander Bhatnagar
mukesh@physics.iitd.ac.in

1 Department of Physics, Indian Institute of Technology, New Delhi 110016, India

2 Applied Physics Department, Delhi Technological University, New Delhi 110042, India

3 Centre for Nanoscience & Technology, Prof. Rajendra Singh Institute for Physical Sciences, VBS Purvanchal University, UP-222003 Jaunpur, India

Published online: 07 November 2020

commonly occurs with a dominating fraction in the melt processed PVDF film, is the most thermodynamically stable state at ambient temperature and pressure [4–6]. However, the pizoresponse in PVDF mainly depends upon the extent of formation of β-phase over the other phases due to its better piezoelectric, ferroelectric and pyroelectric properties. The α-phase possesses TGTG′ (T-trans, G-gauche+, G-′gauche), β-phase has all trans (TTTT) and the γ-phase has T_3GT_3G' conformations [7, 8]. Therefore, the formation of β-phase into PVDF polymer is one of the important parameter for enhancement of piezoelectric properties. Normally the β-phase enhancement is done by subsequent stretching and high voltage poling operation [9] but in the present study the enhancement of β-phase has been done by loading appropriate amount of RGO in PVDF matrix. In addition, organic polymers possess small dielectric permittivity, on the other hand ceramics have large permittivity, the incorporation of ceramic in PVDF polymer can provide better flexibility and higher piezo- response. Therefore, polymer-ceramic composites are the best way to produce large dielectric permittivity from ceramics side and lightweight and flexibility from polymer side [10, 11]. Some reports show that incorporation of different kind of nanofillers in pure PVDF matrix enhances the electroactive polar β-phase as well as improve the dielectric properties of PVDF such nanofillers as diisopropylammonium bromide (DIPAB) [12], graphene nanosheets (GNS) [13], ZnO [14, 15], $BaTiO_3$ [16], $NaNbO_3$ [17], ZnO VANRs [18] and $BiFeO_3$ [19]. On the other hand, the lead based nanomaterial zirconate titanate (PZT) has higher piezoelectric coefficient along with higher dielectric constant and dipole moment but due to its toxicity, complex structure and complicated synthesis procedure limit the use of PZT as nanofiller in PVDF matrix [18, 19].

In this research work, we have synthesized reduced graphene oxide (RGO) nanosheets by sonication assisted method and fabricated a PVDF/RGO nanocomposite thin film via solution casting method. The present study the effect of incorporation of RGO nanosheets in the PVDF matrix has been investigated. We have also investigated the effect of RGO nanosheets on the nucleation of the electroactive polar β-phase, optical, structural and electrical properties of nanocomposite films. Further, we have also measured piezo- response (output voltage signal) when the ultraviolet–visible light has been exposed on the nanocomposite films by applied the pressure from human's finger.

Experimental

Materials

Poly (vinylidene fluoride) (PVDF) pellets with a molecular weight of 530,000 were obtained from Sigma Aldrich, USA. Dry N, N-dimethylformamide (DMF, Merck, India), Graphite powder (100 micron) was procured from SDFCL, Mumbai. Sulfuric acid and orthophosphoric acid and ascorbic acid were purchased from Thermo Fisher Scientific, India.

Synthesis of RGO

Reduced graphene oxide (RGO) was synthesized following a previously reported sonication assisted method [22] with slight modifications. 1 g of graphite powder was added to a 60 mL of acid mixture (concentrated sulfuric acid and orthophosphoric acid at 9:1 v/v ratio) maintained in an ice bath. The mixture was stirred for 10 min following which 3 g $KMnO_4$ was gradually added to this suspension. After 30 min of stirring the reaction mixture was removed from the ice bath and stirred at room temperature for about half an hour followed by 5 min of sonication in an ultrasonic water bath. The process of stirring followed by sonication (5 min each) was repeated for about 4 h following which the reaction was quenched by addition of 200 mL distilled water. After another 2 h of ultrasonication the suspension was raised to a pH of ~ 6 by the addition of 1 (M) NaOH solution followed by another 1 h of ultrasonic treatment. 100 mL 10% ascorbic acid aqueous solution was added to the suspension slowly at room temperature. The suspension was heated at 95 °C for 1 h to carry out the reduction reaction. The resultant black precipitate was filtered out and washed by 1 (N) HCl solution and distilled water till neutral pH was attained followed by drying at room temperature to obtain the black RGO powder.

Synthesis of flexible PVDF/RGO nanocomposite films

The preparation of flexible nanocomposite films of PVDF/RGO depicted in Fig. 1. 0.5 g of PVDF was dissolved in 5 ml of N, N-dimethylformamide (DMF) and stirring with the help of magnetic stirrer at 60 °C for 3 h. Different weight percentage (0, 1, 3, 7 and 11wt%) of RGO sheets were mixed in the above solution and these mixed solutions are named as PVDF, PVDF/RGO1, PVDF/RGO3, PVDF/RGO7 and PVDF/RGO11, respectively. These solutions are then stirred for 1 h and later ultra-sonicated for 2 h to get well dispersed solutions. The well dispersed solution of PVDF/RGO is poured over a clean glass substrate (size 4 cm × 2.5 cm) and kept in oven maintained at 80 °C for 5 h to evaporate the solvent inside an oven to yield the nanocomposite films and then films cool down to room temperature. To obtain free standing nanocomposite films we have dipped the glass substrate with film into the de-ionized water at room temperature. The nanocomposite films peels off from the glass substrate immediately after dipping into the water. As these nanocomposite films are hydrophobic no water gets stuck on them.

Characterization

The structural analysis and surface morphology of synthesized reduce graphene oxide (RGO) nanosheets have been studied using X-ray diffractometer (XRD, Rigaku Ultima-IV) and Transmission Electron Microscope (TEM, JEOL, JEM-2200-FS) respectively. The characteristic modes of RGO have been analyzed by Micro Raman Spectroscopy (MRS). The structural properties of RGO loaded PVDF nanocomposite films have been studied by XRD. The surface morphology of the nanocomposite films has been observed with a Scanning Electron Microscope (SEM, Hitachi, TM3000). The formation of β-phase in the nanocomposite films has been characterized by Fourier Transform Infra-Red (FTIR) spectroscopy (NICOLET 5700). The dielectric property has been investigated by using LCR meter (HIOKI, IM3536) in the frequency range 4 to 10^6 Hz at room temperature. The optical properties are carried out by ultraviolet visible absorption spectroscopy (UV–VIS PerkinElmer). The polarization–electric field (P-E) loop of PVDF and PVDF/RGO nanocomposite films has been studied by using a P-E loop tracer. The piezo-response, in terms of output voltage, has been measured using by digital storage oscilloscope (DSO) under the finger tapping mode from a fixed height.

Results and discussions

Characterization of Reduce Graphene Oxide (RGO) sheets

Figure 2(a) presents the XRD pattern of the RGO powder. The broad peak appearing around 2θ = 22.5° is characteristic of the (002) planes of reduced graphene oxide. The position and the broadness of the peak indicate excellent exfoliation of the graphite sheets used as precursor. A much smaller peak observed at ~ 43° corresponds to the turbostratic band of disordered carbon materials [23].

Figure 2(b) represents the TEM micrograph of the as synthesized RGO nanosheets. The transparent nature indicates well exfoliated feature of the material with only a few layers of RGO to fabricate such sheets. Presence of wrinkles and agglomerated sheets is also well documented feature of oxidation/reduction method.

Raman spectroscopy has been known to be an effective technique for identifying structure, defects, disorder and number of graphitic layers in carbon materials. Figure 2(c) represents the Raman spectra of the as synthesized RGO and a lower intensity of G band at 1573 cm^{-1} and higher intensity of D band at 1340 cm^{-1} were observed. The D-band is associated with defects, wrinkles and dangling bonds in the carbon structure while the G-band whose intensity is proportional to the number of graphitic layers, is associated with the in-plane vibration of the sp^2 bonded carbon atoms. I_D/I_G was found to be 1.19 which is indicative of the presence of extensive number of defects in the carbon structure as a result of continuous chemical oxidation/reduction treatment whereas lowered G band intensity was indicative of decreased number of graphitic layers. A relatively weak 2D peak at 2670 cm^{-1} and a S3 band at 2910 were also observed. The slight downshifting of the 2D peak indicates the presence of very few graphitic layers within the nanosheets. The S3 band is a characteristic second order peak due to D-G recombination.

Characterization of PVDF/RGO nanocomposite films

XRD analysis

To reveal the crystal phase of PVDF/RGO nanocomposite films, X-ray diffraction (XRD) measurement has been carried out and results are shown in Fig. 3. Figure 3 shows the 2θ peaks for pure PVDF at 18.21° and 20.22° corresponding to planes (020), reflection of α-phase and (110), reflection of β-phase respectively [24]. The presence of β-phase in pure PVDF is due to dipole moment of DMF [25], therefore, the spectra of pure

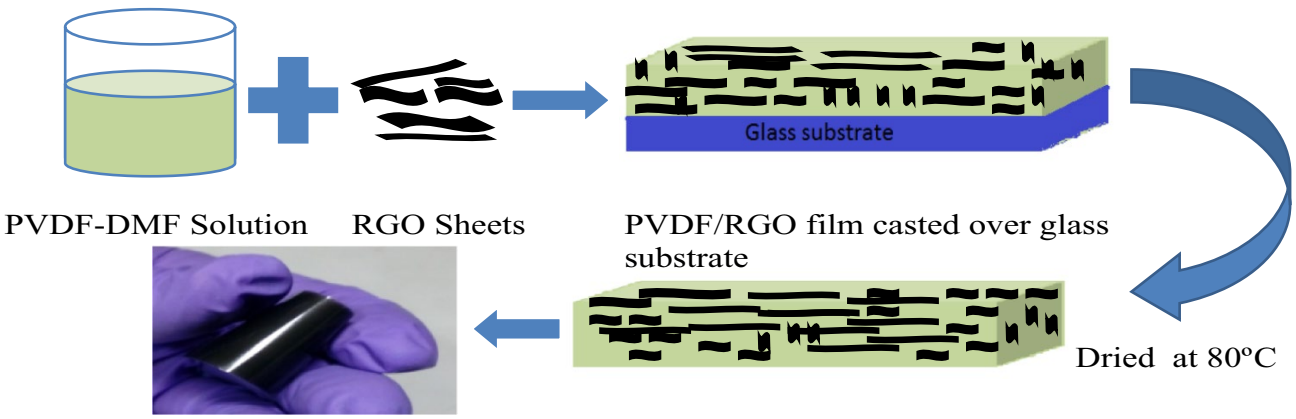


Fig. 1 Schematic diagram representation for the synthesis of flexible PVDF/RGO nanocomposite films

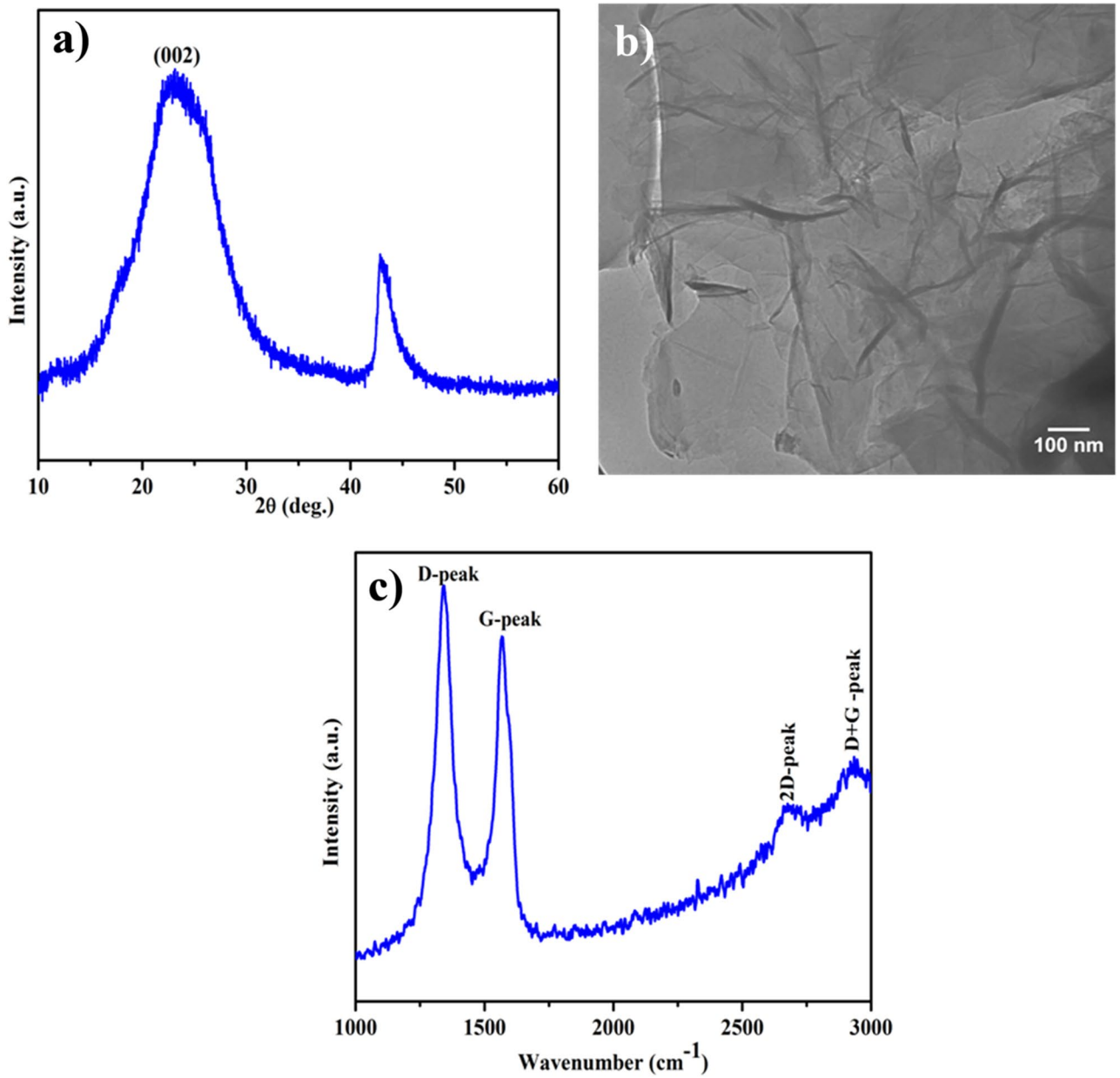


Fig. 2 (**a**) XRD pattern of RGO sheets, (**b**) TEM image of RGO sheets and (**c**) Raman spectra of RGO

PVDF exhibit both α- and β-phases. Figure 3 shows that as the concentration of RGO increases in pure PVDF, the peak intensity corresponding to β-phase increases as compared to pure PVDF, but α-phase still exists in the nanocomposite films.

It indicates that the presence of small amount of RGO in pure PVDF is sufficient to nucleate electroactive β-phase in PVDF polymer. As shown in Fig. 3, while the concentration of nanofillers is increases, the peak intensity corresponding to β-phase is continuously increasing, which enhances the overall piezoelectric response of PVDF polymer.

SEM analysis

Surface morphology of nanocomposite films is shown in Fig. 4a-c. The pure PVDF films show uniform smooth surface as shown in Fig. 4a. Figure 4 (b, c) shows randomly distributed RGO nanosheets in the PVDF matrix.

FTIR Spectra of PVDF/RGO nanocomposite films

Fourier transform infra-red (FTIR) spectroscopy has been used for analyzing the phase formation of nanocomposite PVDF films having RGO weight percentages of 0, 1, 3, 7

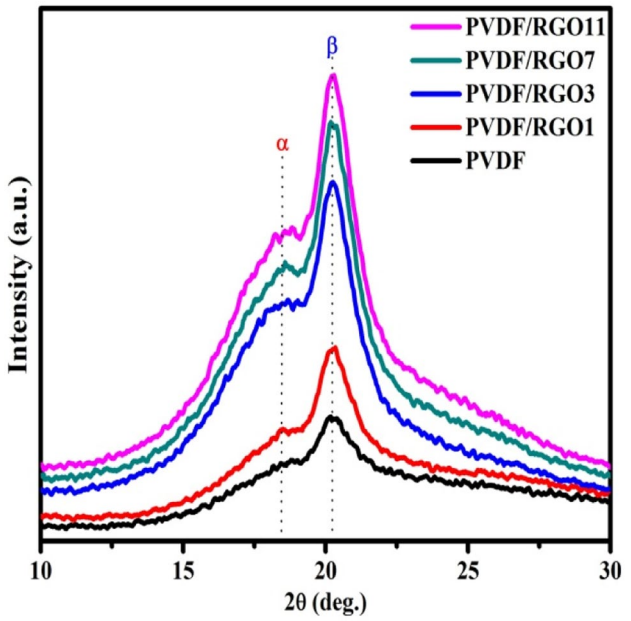


Fig. 3 X-ray diffraction patterns (XRD) of pure PVDF and PVDF/RGO nanocomposite at different RGO loading

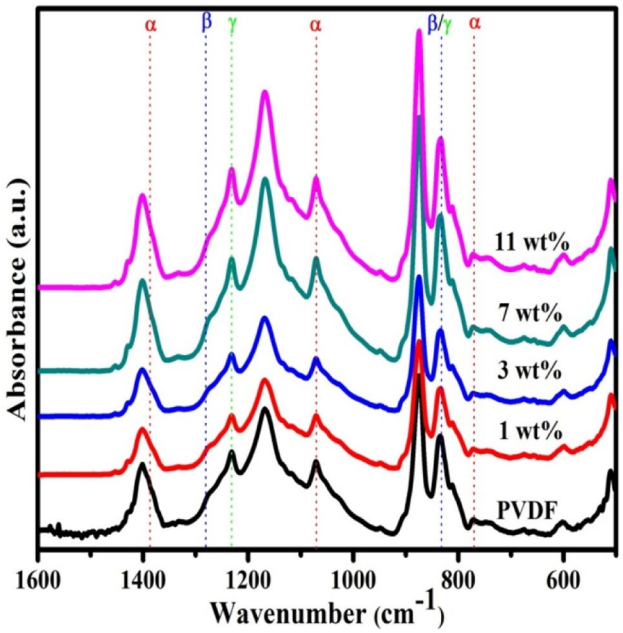


Fig. 5 FTIR spectra of pure PVDF and PVDF/RGO nanocomposite films at different RGO loading

and 11 wt%. The results are shown in Fig. 5. The characteristics peaks of absorption band of the non-polar (α) and (β), (γ) polar phases of PVDF/RGO nanocomposite are identified at 408, 510, 765, 840, 1210, 1232, 1280 and 1384 cm^{-1} respectively. The peaks corresponding to the non-polar (α) phase are 408, 765, 1210 and 1384, polar (β) phase are 510, 840, and 1280, and polar (γ) phase are 840 and 1232 [20, 21, 28]. Some reports show that the absorption bands for β and γ phases of nanocomposite films, superimposed due to common (TTTT) confirmation, that creates difficulty for the phases separation [22, 23].

To identified the relative amount of β-phase F(β) present in the samples is calculated using the Beer-Lambert law [27].

$$F(\beta) = \frac{X_{(\beta)}}{X_{(\alpha)} + X_{(\beta)}} \times 100\% = \frac{A_{(\beta)}}{\left(\frac{k_\beta}{k_\alpha}\right)A_{(\alpha)} + A_{(\beta)}} \times 100\% \quad (1)$$

Where, A_α and A_β represent the absorption intensity of the α-phase and β-phase (at 764 and 840 cm^{-1}), respectively. X_α and X_β are the corresponding crystallinities of the α-phase and β-phase; k_α and k_β are corresponding absorption coefficients with the values $6.1 \times 10^4 cm^2/mol$ and $7.7 \times 10^4 cm^2/mol$ respectively.

Figure 5 shows that the relative amount of F(β) for the PVDF film is 53% and for PVDF/RGO1, PVDF/RGO3, PVDF/RGO7 and PVDF/RGO11 nanocomposite films are 61, 64.5, 69.2 and 70%, respectively. A slightly improvement

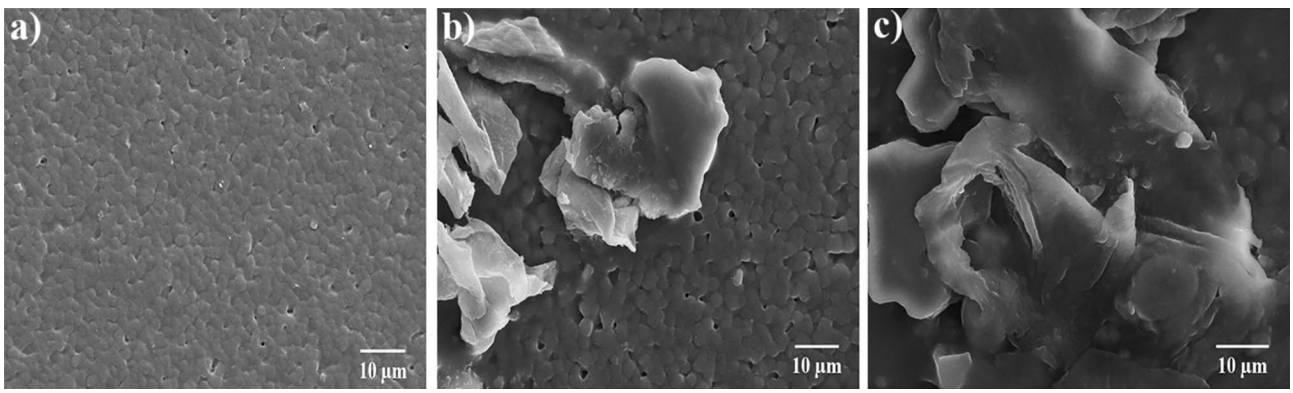


Fig. 4 Scanning electron microscope (SEM) images of (**a**) polyvinylidene fluoride (PVDF), (**b**) PVDF/RGO3 and (**c**) PVDF/RGO11

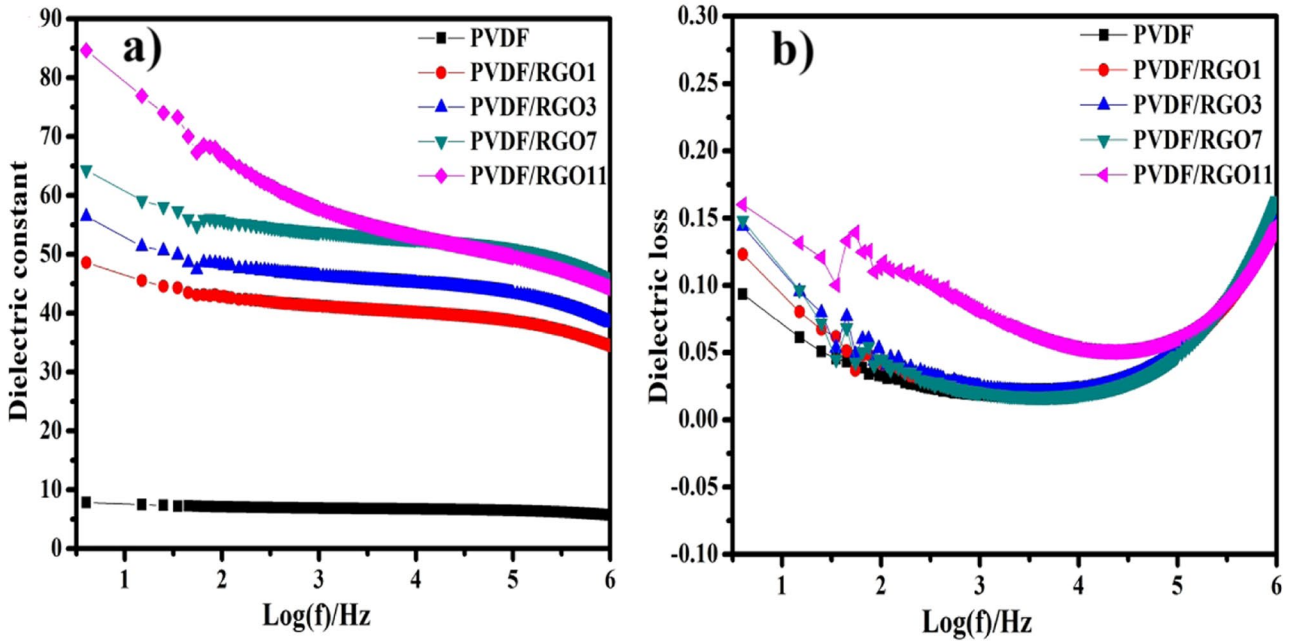


Fig. 6 The frequency dependent (**a**) dielectric constant and (**b**) dielectric loss of PVDF/RGO nanocomposite films

of β-phase in PVDF has been observed after addition of RGO in PVDF polymer. The result indicates that RGO play an important role in the nucleation of the β-phase. FTIR analysis is in good agreement with XRD analysis.

Electrical properties of PVDF/RGO nanocomposite films

Figure 6(a, b) show variation of dielectric constant (ε') (Fig. 6a) and dielectric loss (ε'') (Fig. 6b) of different PVDF/RGO nanocomposite films as a function of frequency at room temperature.

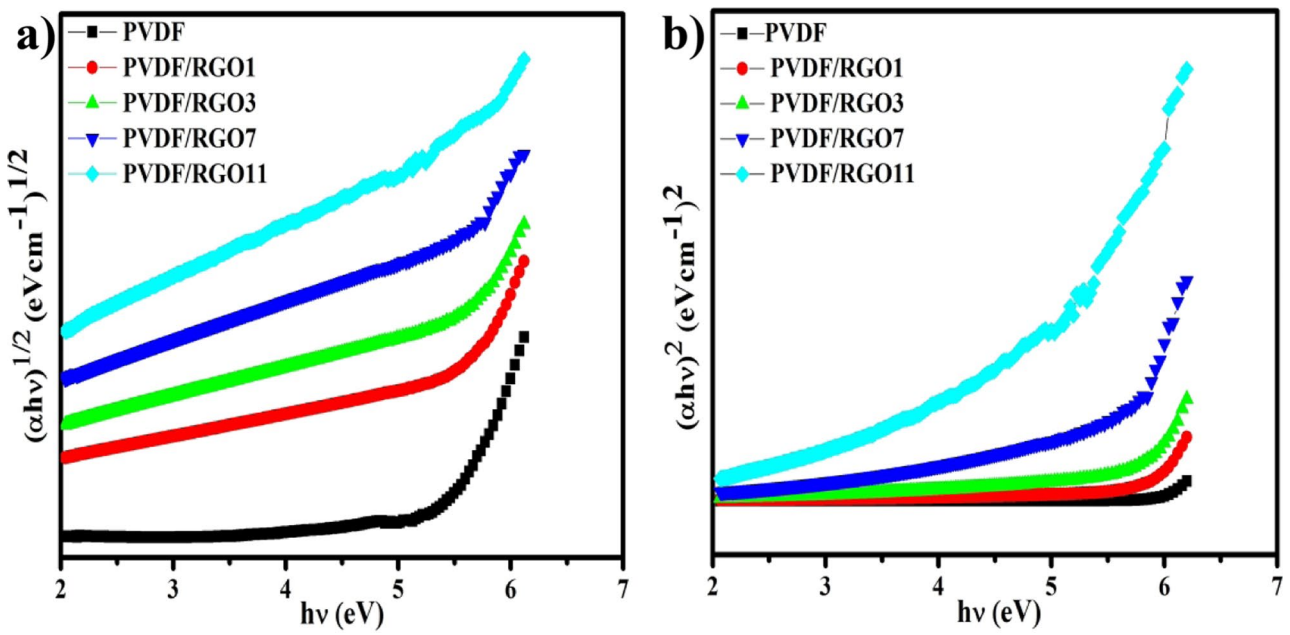


Fig. 7 Tauc plot of (**a**) Indirect band gap and (**b**) Direct band gap of PVDF and PVDF/RGO nanocomposite films

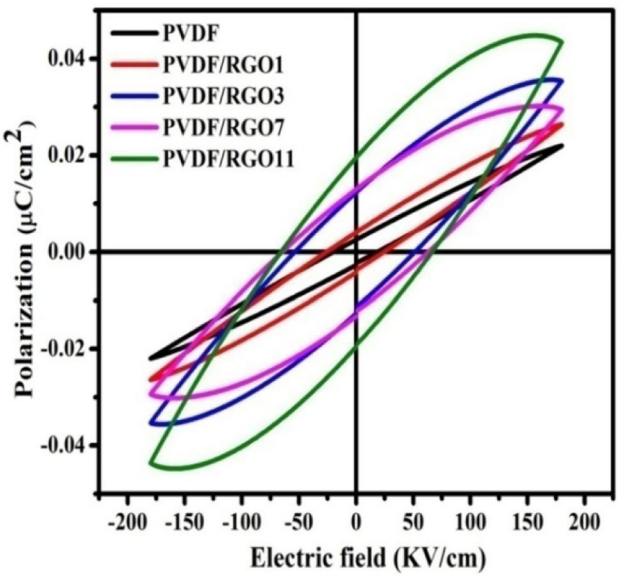


Fig. 8 Hysteresis P-E loops for pure PVDF, PVDF/RGO1, PVDF/RGO3, PVDF/RGO7 and PVDF/RGO11 nanocomposite films

Figure 6(a) shows the remarkable improvement in dielectric constant and dielectric loss even at addition of 0.1wt% of RGO in neat PVDF. The value of dielectric constant (ε') and dielectric loss (ε'') for neat PVDF at 1 kHz are 6.8 and 0.0184, but the corresponding value of dielectric constant (ε') and dielectric loss (ε'') for 0.1wt% of RGO at the same frequency are 42 and 0.023. The dielectric constant (ε') for 0.1wt% of RGO is about six times higher than that of pure PVDF. This characteristic has been explained on the basis of Maxwell-Wager-Sillars (MWS) polarization mechanism and space charge polarization [3, 24]. When an electric field is applied to nanocomposite films, the charges accumulated at interface with in the bulk sample and interface between sample and the electrodes.

This charges carrier accumulation at the interface induces interfacial polarization and ultimately enhancement in dielectric constant (ε'), because of two heterogeneous system consisting different dielectric materials. The large interfacial areas in composite provide numerous sites for the reinforced MWS effect. This phenomenon has appeared at the lower frequency because it took times to complete MWS polarization. On the other hand, Space charge polarization happened by non-uniformly space charge in heterogeneous nanocomposite [32]. Both mechanisms are responsible to enhancement of dielectric constant (ε') and dielectric loss (ε'') in polymer based nanocomposites. The values of dielectric constant (ε') and dielectric loss (ε'') increased slowly with the loading of RGO content in PVDF. The value of dielectric loss (ε'') for PVDF and PVDF/RGO4 nanocomposite films, as shown in Fig. 6(b), is respectively 0.0184 and 0.0805 at 1 kHz which is nearly 4.5 times.

The Optical band gap

The optical properties of PVDF and PVDF/RGO nanocomposite films were investigated using UV–Visible spectroscopy over the range of 200-900 nm. Figure 7 shows that tauc plot of indirect and direct band gap of PVDF and PVDF/RGO nanocomposite films using this equation [33].

$$\alpha(\upsilon) = 2.303\left(\frac{A}{t}\right) \tag{2}$$

Where A is the absorbance (a.u.) and t is the thickness of nanocomposite films.

$$\alpha(\upsilon) = \frac{(h\upsilon - E_g)^n}{h\upsilon} \tag{3}$$

Where the exponent n=2 for allowed direct transition, while $n = \frac{1}{2}$ for indirect transition. Therefore, indirect band gaps and direct band gap of PVDF/RGO nanocomposite films are determined by Tauc plots. Figure 7 (a) shows the plot of $(\alpha h\upsilon)^{\frac{1}{2}}$ versus photon energy ($h\upsilon$) for PVDF/RGO nanocomposite films. The value of optical band gap energy is evaluated from the intersection of extrapolated line with the photon energy axis at $\alpha=0$. The values of indirect band gap for PVDF/RGO nanocomposites are 5.31, 4.74, 4.46, 3.93 and 3.58 eV. Therefore, the incorporation of RGO nanosheets in PVDF polymer, leads to decreases the optical band gap due to formation of defects and states near the band edges, lowering the fermi level in nanocomposite films.

Direct band gaps are obtained from the plots of $(\alpha h\upsilon)^2$ versus photon energy ($h\upsilon$) as shown in Fig. 7 (b). The values of direct band gap for PVDF/RGO nanocomposites are 5.62, 5.44, 5.23, 4.85 and 4.42 eV. This decrease in optical band gap values are due to the formation of defects/ or interaction between π -electron of RGO with CH_2- dipoles of PVDF after addition of RGO nanosheets in PVDF polymer [30, 31]. The optical energy band gaps for indirect and direct transition for this system both decreases as the RGO (wt%) concentration is increases. This decrease in the optical energy band gaps is due to formation of defects, and these defects are formed lower energy states and responsible for increases the number of charge carriers in the conduction band [35]. These defects are also helps to find continues chain to travel the charge carries resulting in delocalized π-electron of RGO. This change indicates a lowering in energy band gap leads to an increase in the electrical conductivity of the polymer and consequently influences the optical properties nanocomposite films. Hence the optical band gap is found to be compositional and structural dependence.

Ferroelectric properties

To confirm the enhancement of ferroelectric properties of PVDF/RGO nanocomposite films, P-E loop measurement has been carried out. Figure 8 represents the hysteresis loop

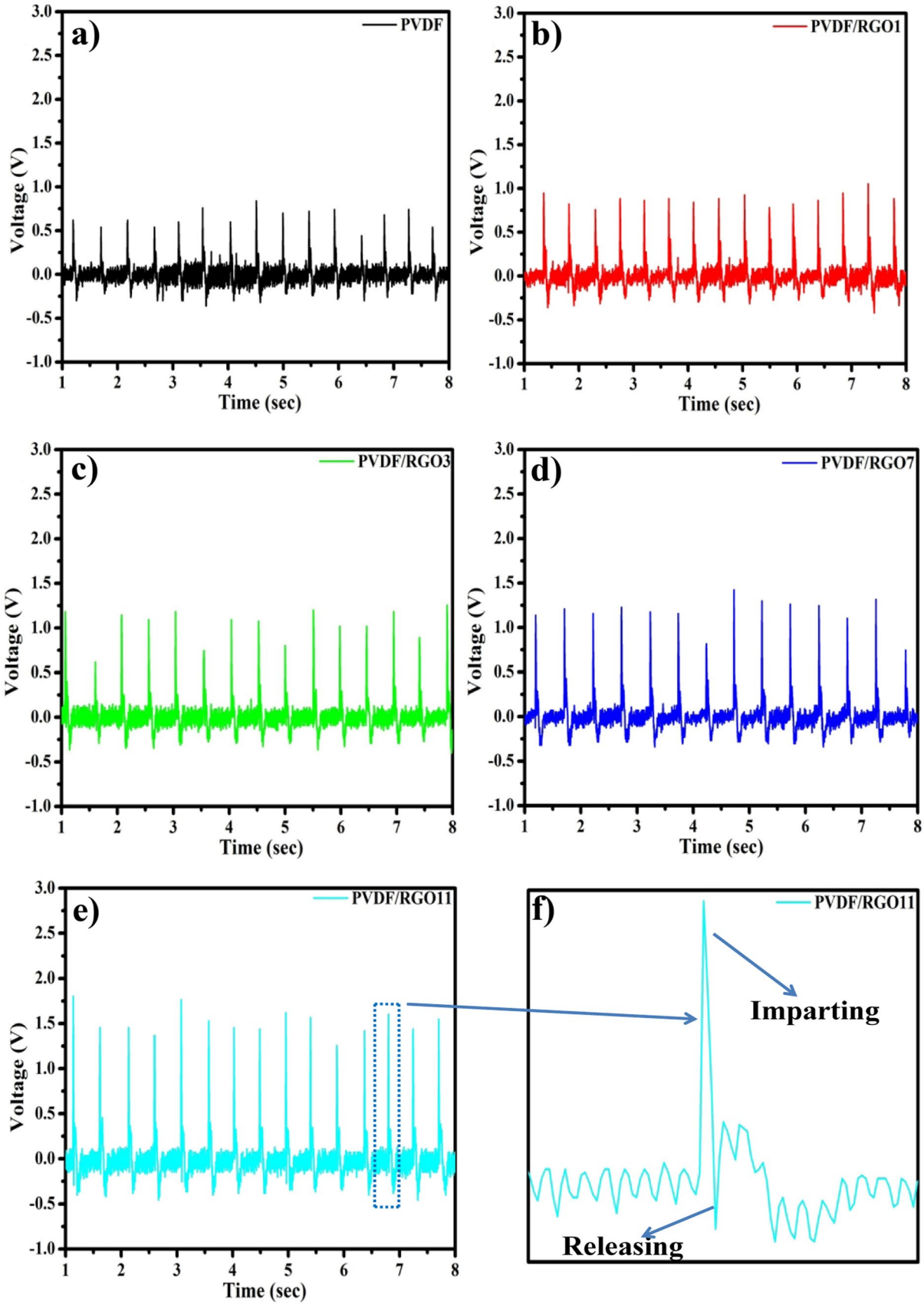


Fig. 9 Output voltage response (**a**) PVDF, (**b**) PVDF/RGO1, (**c**) PVDF/RGO3, (**d**) PVDF/RGO7, (**e**) PVDF/RGO11 and (**f**) Touch and release response from nanocomposite films

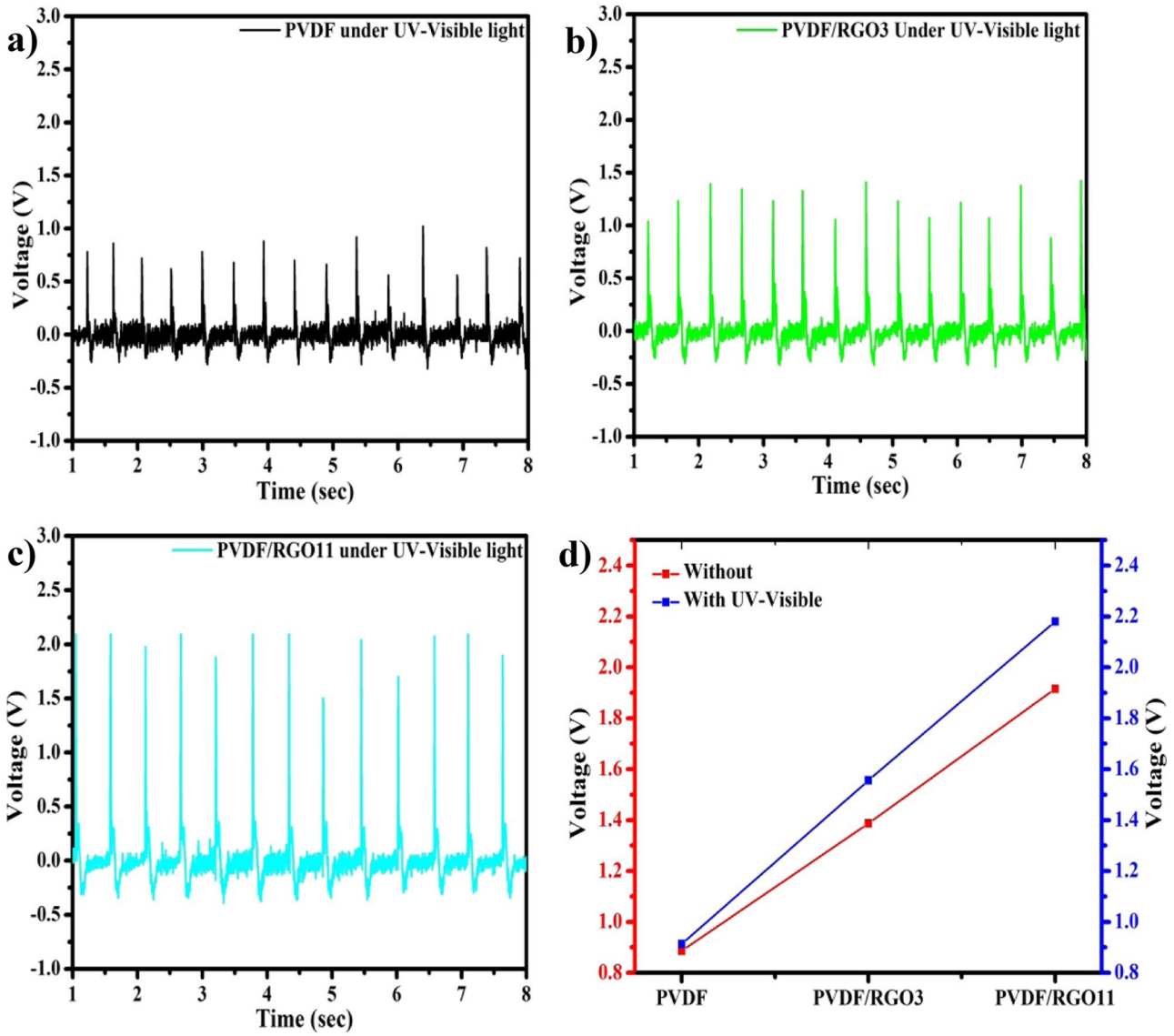


Fig. 10 Output voltage response under ultraviolet–visible light (**a**) PVDF, (**b**) PVDF/RGO3, (**c**) PVDF/RGO11 and (**d**) Comparative output voltage of PVDF/RGO nanocomposite films without and with UV Vis light

without electrical poling. The area of upper electrode was kept at 7.25cm^2 whereas thickness of nanocomposite was kept at 20μm for all the samples. The values of remnant polarization (P_r) are 0.0025, 0.0040, 0.0122, 0.0129 and 0.01946μC/cm^2 for PVDF, PVDF/RGO1, PVDF/RGO3, PVDF/RGO7 and PVDF/RGO11 nanocomposite films respectively under the applied electric field at 180 kV/cm. From the above results, it can be concluded that the with increment of RGO concentration in PVDF matrix, the value of remnant polarization also increases due to hetro polarization [36]. The maximum value of remnant polarization (P_r) is found to be 0.01946μC/cm^2 for PVDF/RGO11 nanocomposite film.

When the RGO sheets loading were at 11 wt%, the value of remnant polarization (P_r) is about seven times higher than that of pure PVDF. This result proves that the loading of RGO nanosheets in PVDF matrix to enhance piezoelectric properties of nanocomposite films. Therefore, the effect of RGO sheets on the nucleation of β-phase in PVDF matrix mainly depends on the (i) concentration (ii) interaction between π-electron with CH_2-dipoles and (iii) oxygen functionalized group on the surface of RGO that can interact with CF_2- dipoles in PVDF via electrostatic interaction and (iv) processing condition [22, 31, 33].

Piezoelectric energy harvesting performance of the nanocomposite films

To check the piezoresponse and energy harvesting performance of the PVDF/RGO nanocomposite films, using

electrode–composite-electrode (ECE) and to fabricate flexible nanogenerator, a 60 nm thin film of aluminum has been deposited on the both sides of film by thermal evaporation technique. The active area on both sides of film is (3.20 cm × 2.25 cm). The piezoresponse and energy harvesting capability from the PVDF/RGO nanocomposite were measured without applying electrical poling.

Figure 9 shows that the output voltage response from PVDF/RGO nanocomposite films, when the application of mechanical pressure (10-12 kPa) was applied by human finger pressure imparted on the surface of PVDF/RGO nanocomposite films. As shows in Fig. 9 the average of peak to peak output voltages are 0.886, 1.234, 1.387, 1.523, and 1.915 V for PVDF, PVDF/RGO1, PVDF/RGO3, PVDF/RGO7 and PVDF/RGO11 nanocomposite films respectively. This enhancement of output voltage at the same pressure is due to interaction between π-electrons of RGO with CH_2- dipoles of PVDF [21, 37]. The positive peak is obtained for imparting and negative peak obtained when the pressure is releasing as shown in Fig. 9(f).As the concentration of RGO increased in PVDF matrix, RGO provides a conducting paths for charges induced inside the nanocomoposite to move the surface of PVDF/RGO nanocomposite films and further allows the dipoles of PVDF to align [6].

Energy harvesting performance of nanocomposite films under ultraviolet–visible light

When ultraviolet–visible light of wavelength (300-900 nm) was exposed on nanocomposite films, the average output peak to peak voltages were observed as 0.913, 1.556 and 2.180 V for pure PVDF, PVDF/RGO3 and PVDF/RGO11 nanocomposite films respectively after applying same pressure by finger as mentioned above. The results are shown in Fig. 10. There is small increment in voltage for PVDF nanocompoiste film under the UV Vis light. But when the RGO nanofiller incorporated into PVDF matrix and exposed UV Vis light on the nanocomposite, the changes of output voltages shows in Fig. 10, these enhancements of output voltage under the UV Vis light at the same pressure have enough energy provide to nanocomposite films and these energy has to help the alignment of dipole in PVDF definite direction.

The nanocomposite (PVDF/RGO11) generates maximum open circuit voltage of 2.180 V with human finger under UV Vis light which is about 13.8% higher as compared to the case of unexposed UV Vis exposure.

Conclusion

Nanocomposite films of pure PVDF and PVDF/RGO were made via solution casting method and their piezo-properties have been studied. FTIR analysis was used to determine the content of β-phase in different RGO loaded PVDF nanocomposites**.** The hysteresis loop indicates the charge storage capability of the material. The value of the remnant polarization is directly proportional to piezo-response of material. The voltage signal found to be maximum for PVDF/RGO11 nanocomposite as compared to pure PVDF. The maximum value of the voltage signal is 1.915 V under the UV Vis light exposure which was 13.8% higher as compared to unexposed sample. The results also indicate that after addition of RGO and applying UV vis exposure, the piezo response can be enhanced by more than 8 times as compared to pure PVDF films. The enhancement can be attributed to (i) good interaction between of π-electron RGO with CH_2- dipoles of PVDF and (ii) RGO provides the conducting path for charges to move to surfaces. Both these mechanisms contribute to overall piezo-response. Hence it can be concluded that large area, flexible nanocomposite films with very high piezo-response can be fabricated by incorporation of RGO nanosheets in PVDF polymer.

Acknowledgments The author would like to acknowledge the University Grant Commission (UGC), Government of India, for providing the research fellowship.

References

1. Jana S, Garain S, Ghosh SK, Sen S, Mandal D (2016) The preparation of γ-crystalline non-electrically poled photoluminescant ZnO-PVDF nanocomposite film for wearable nanogenerators. Nanotechnology 27:1–12. https://doi.org/10.1088/0957-4484/27/44/445403
2. Heywang W, Lubitz K, Wersing W (2008) Piezoelectricity: Evolution and Future of a Technology, Springer Berlin Heidelberg. https://books.google.co.in/books?id=8KyawZ92eIEC
3. Hu YC, Hsu WL, Wang YT, Ho CT, Chang PZ (2014) Enhance the pyroelectricity of polyvinylidene fluoride by grapheneoxide doping. Sensors (Switzerland) 14:6877–6890. https://doi.org/10.3390/s140406877
4. Park JH, Kurra N, AlMadhoun MN, Odeh IN, Alshareef HN (2015) A two-step annealing process for enhancing the ferroelectric properties of poly(vinylidene fluoride) (PVDF) devices. Journal of Materials Chemistry C 3:2366–2370. https://doi.org/10.1039/c4tc02079k
5. Kar E, Bose N, Das S, Mukherjee N, Mukherjee S (2015) Enhancement of electroactive β phase crystallization and dielectric constant of PVDF by incorporating GeO2and SiO2nanoparticles. Physical Chemistry Chemical Physics 17:22784–22798. https://doi.org/10.1039/c5cp03975d
6. Singh HH, Singh S, Khare N (2017) Design of flexible PVDF/NaNbO3/RGO nanogenerator and understanding the role of nanofillers in the output voltage signal. Compos Sci Technol 149:127–133. https://doi.org/10.1016/j.compscitech.2017.06.013
7. Lu Y, Claude J, Neese B, Zhang Q, Wang Q (2006) A modular approach to ferroelectric polymers with chemically tunable Curie temperatures and dielectric constants. J Am Chem Soc 128:8120–8121. https://doi.org/10.1021/ja062306x
8. Teka A, Bairagi S, Shahadat M, Joshi M, Ziauddin Ahammad S, Wazed Ali S (2018) Poly(vinylidene fluoride) (PVDF)/potassium sodium niobate (KNN)-based nanofibrous web: A unique nanogenerator for renewable energy harvesting and investigating

the role of KNN nanorods. Polym Adv Technol 29(9):2537–2544. https://doi.org/10.1002/pat.4365
9. Huang L, Lu C, Wang F, Wang L (2014) Preparation of PVDF/graphene ferroelectric composite films by in situ reduction with hydrobromic acids and their properties. RSC Advances 4:45220–45229. https://doi.org/10.1039/c4ra07379g
10. Kim HS, Kim JH, Kim J (2011) A review of piezoelectric energy harvesting based on vibration. International Journal of Precision Engineering and Manufacturing 12:1129–1141. https://doi.org/10.1007/s12541-011-0151-3
11. Niu Y, Yu K, Bai Y, Xiang F, Wang H (2015) Fluorocarboxylic acid-modified barium titanate/poly(vinylidene fluoride) composite with significantly enhanced breakdown strength and high energy density. RSC Advances 5:64596–64603. https://doi.org/10.1039/c5ra09023g
12. Sui Y, Chen WT, Ma JJ, Hu RH, Liu DS (2016) Enhanced dielectric and ferroelectric properties in PVDF composite flexible films through doping with diisopropylammonium bromide. RSC Advances 6:7364–7369. https://doi.org/10.1039/c5ra25371c
13. Shang J, Zhang Y, Yu L, Shen B, Lv F, Chu PK (2012) Fabrication and dielectric properties of oriented polyvinylidene fluoride nanocomposites incorporated with graphene nanosheets. Mater Chem Phys 134:867–874. https://doi.org/10.1016/j.matchemphys.2012.03.082
14. Gaur MS, Indolia AP (2011) Thermally stimulated dielectric properties of polyvinylidenefluoride-zinc oxide nanocomposites. J Therm Anal Calorim 103:977–985. https://doi.org/10.1007/s10973-010-1190-1
15. Shin KY, Lee JS, Jang J (2016) Highly sensitive, wearable and wireless pressure sensor using free-standing ZnO nanoneedle/PVDF hybrid thin film for heart rate monitoring. Nano Energy 22:95–104. https://doi.org/10.1016/j.nanoen.2016.02.012
16. Yaqoob U, Uddin ASMI, Chung G-S (2016) The effect of reduced graphene oxide on the dielectric and ferroelectric properties of PVDF–$BaTiO_3$ nanocomposites. RSC Advances 6:30747–30754. https://doi.org/10.1039/C6RA03155B
17. Anand A, Bhatnagar MC (2019a) Effect of sodium niobate (NaNbO 3) nanorods on β-phase enhancement in Polyvinylidene fluoride (PVDF) polymer Effect of sodium niobate (NaNbO 3) nanorods on β-phase enhancement in Polyvinylidene fluoride (PVDF) polymer, Material Research Express
18. Anand A, Bhatnagar MC (2019b) Role of vertically aligned and randomly placed zinc oxide (ZnO) nanorods in PVDF matrix : Used for energy harvesting. Materials Today Energy 13:293–301. https://doi.org/10.1016/j.mtener.2019.06.005
19. Bhadra D, Masud MG, Sarkar S, Sannigrahi J, De SK, Chaudhuri BK (2012) Synthesis of PVDF/BiFeO3 nanocomposite and observation of enhanced electrical conductivity and low-loss dielectric permittivity at percolation threshold. J Polym Sci, Part B: Polym Phys 50:572–579. https://doi.org/10.1002/polb.23041
20. Dietze M, Es-Souni M (2008) Structural and functional properties of screen-printed PZT-PVDF-TrFE composites. Sens Actuators, A 143:329–334. https://doi.org/10.1016/j.sna.2007.11.016
21. Karan SK, Mandal D, Khatua BB (2015) Self-powered flexible Fe-doped RGO/PVDF nanocomposite: An excellent material for a piezoelectric energy harvester. Nanoscale 7:10655–10666. https://doi.org/10.1039/c5nr02067k
22. Abdolhosseinzadeh S, Asgharzadeh H, Kim HS (2015) Fast and fully-scalable synthesis of reduced graphene oxide. Scientific Reports 5:1–7. https://doi.org/10.1038/srep10160
23. Dave K, Park KH, Dhayal M (2015) Two-step process for programmable removal of oxygen functionalities of graphene oxide: functional, structural and electrical characteristics. RSC Advances 5:95657–95665. https://doi.org/10.1039/c5ra18880f
24. El Achaby M, Arrakhiz FZ, Vaudreuil S, Essassi EM, Qaiss A (2012) Applied Surface Science Piezoelectric [illegible] -polymorph formation and properties enhancement in graphene oxide – PVDF nanocomposite films. Appl Surf Sci 258:7668–7677. https://doi.org/10.1016/j.apsusc.2012.04.118
25. Benz M, Euler WB, Gregory OJ (2002) The role of solution phase water on the deposition of thin films of poly(vinylidene fluoride). Macromolecules 35:2682–2688. https://doi.org/10.1021/ma011744f
26. Kanik M, Aktas O, Sen HS, Durgun E, Bayindir M (2014) Spontaneous High Piezoelectricity in Poly(vinylidene fluoride) Nanoribbons Produced by Iterative Thermal Size Reduction Technique. ACS Nano 8:9311–9323. https://doi.org/10.1021/nn503269b
27. Martins P, Lopes AC, Lanceros-Mendez S (2014) Electroactive phases of poly(vinylidene fluoride): Determination, processing and applications. Prog Polym Sci 39:683–706. https://doi.org/10.1016/j.progpolymsci.2013.07.006
28. Pan CT, Kun Yen C, Wu HC, Lin L, Lu Y-S, Huang JC, Kuo SW (2015) Significant piezoelctric and energy hrvesting enhancment of poly(vinyledene fluoride)/polypeptide fiber composites prepared through near-field electrospinning. Journal of Materials Chemistry A 3:6835–6843. https://doi.org/10.1039/C5TA00147A
29. Lopes AC, Costa CM, Tavares CJ, Neves IC, Lanceros-Mendez S (2011) Nucleation of the electroactive γ phase and enhancement of the optical transparency in low filler content poly(vinylidene)/clay nanocomposites. J Phys Chem C 115:18076–18082. https://doi.org/10.1021/jp204513w
30. Biswas A, Henkel K, Schmeißer D, Mandal D (2017) Comparison of the thermal stability of the α, β and γ phases in poly(vinylidene fluoride) based on *in situ* thermal Fourier transform infrared spectroscopy. Phase Transitions 1594:1–9. https://doi.org/10.1080/01411594.2017.1337902
31. Lu H, Zhang X, Zhang H (2006) Influence of the relaxation of Maxwell-Wagner-Sillars polarization and dc conductivity on the dielectric behaviors of nylon 1010. J Appl Phys 100:054104. https://doi.org/10.1063/1.2336494
32. Zha JW, Dang ZM, Song HT, Yin Y, Chen G (2010) Dielectric properties and effect of electrical aging on space charge accumulation in polyimide/TiO_2 nanocomposite films. J Appl Phys 108:94113–94116. https://doi.org/10.1063/1.3506715
33. Indolia AP, Gaur MS (2013) Optical properties of solution grown PVDF-ZnO nanocomposite thin films. J Polym Res 20:43. https://doi.org/10.1007/s10965-012-0043-y
34. Kilarkaje S, Manjunatha V, Raghu S, Ambika Prasad MVN, Devendrappa H (2011) Effect of 8 MeV electron irradiation on the optical properties of doped polymer electrolyte films. J Phys D Appl Phys 44:105403. https://doi.org/10.1088/0022-3727/44/10/105403
35. Lakshmi GBVS, Ali V, Siddiqui AM, Kulriya PK, Zulfequar M (2009) Studies on structural, optical and cluster size of poly(m-toluidine)- polyvinyl chloride blends. Radiat Eff Defects Solids 164:162–169. https://doi.org/10.1080/10420150902764186
36. Rahman A, Lee B, Phan D (2013) Fabrication and characterization of highly efficient flexible energy harvesters using PVDF–graphene nanocomposites. Smart Mater Struct 22:085017–085027. https://doi.org/10.1088/0964-1726/22/8/085017
37. Yang L, Qiu J, Ji H, Zhu K, Wang J (2014) Composites : Part A Enhanced dielectric and ferroelectric properties induced by TiO 2 @ MWCNTs nanoparticles in flexible poly (vinylidene fluoride) composites. Compos A 65:125–134. https://doi.org/10.1016/j.compositesa.2014.06.006
38. Anand A, Meena D, Bhatnagar MC (2020) Synthesis and characterization of fl exible PVDF/Bi2Al4O9/RGO based piezoelectric materials for nanogenerator application. J Alloy Compd 843:156019. https://doi.org/10.1016/j.jallcom.2020.156019

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Experimental Techniques
https://doi.org/10.1007/s40799-020-00411-7

RESEARCH PAPER

Estimation and Experimental Validation of Mean-Field Homogenised Effective Properties of Composite

A. Sharma[1] · V. Rastogi[1] · A.K. Agrawal[1]

Received: 27 December 2019 / Accepted: 26 October 2020
© The Society for Experimental Mechanics, Inc 2020

Abstract

For composite materials, the system response changes abruptly with a change in the properties of the material. Therefore, attaining significant knowledge about the effect of the material composition on the material properties is crucial. The researchers are looking for new computational methods which can predict these alterations so that the effort in experimental testing can be reduced. In this direction, this paper presents a robust and novel methodology of validating the estimation of the composite's effective through a multi-scale approach by a set of standardized experimentation. These effective properties are estimated through the mean-field homogenization technique whose parameters are driven from the image analysis of Scanning Electron Microscopy (SEM) images. The predicted results are validated with the results obtained by the experimentation as per ASTM E1876 standard. The estimated error between the predicted properties and the experimental values increased with the increase of alumina particle fraction in the matrix. The mean-field homogenization lags behind the experiments for the parameters defined by the image analysis method than the experimental results. The upper bounds of the mean-field homogenization can be used for the composites with higher reinforcement volume fraction.

Keywords Multi-scale analysis · Mean-field homogenization · Image analysis method · ASTM E1876 · Impact hammer test

Introduction

With the advent of composite materials, industries have started focusing on the use of lighter weight materials with the same mechanical properties. Metal matrix composites have the edge over the parent metals for rotor applications, as it has a higher specific modulus, tensile strength and other mechanical properties. Aluminium /alumina MMCs have shown prominent growth in the composite material market because of their compatibility to the rotor systems. [1–3]. Many manufacturers such as Duralcan, G.K.N., Toyota, GM and Nissan have used MMC materials to develop various components with both static and dynamic applications [4].

Metal matrix composites can be manufactured by various methods from liquid and semi solid-state of matrix mixed with solid reinforcement particles. There are various methods such as powder metallurgy, diffusion bonding, infiltration, stir casting and spray forming to manufacture MMC as per requirements. Stir casting is one of the most inexpensive methods which has been used for the fabrication of metal matrix composites. For complicated and extensive size design of metal matrix composites, stir casting is very convenient and useful [5].

A computation tool that provides results validated by experimental or real-time results saves a lot of time and energy for industries. Researchers seek for more tools that can provide accurate and reliable results for metal matrix composites properties. Multi-scale computation of composites is a two-phase process, where the effective macroscopic properties are derived through micro-mechanics. Analysis of composites with discrete reinforcements with random orientation has been found in archival literature [6–9]. There are various approaches through which these effective properties can be predicted efficiently [10]. For composites with a low volume fraction of inclusions, Eshelby's method [11] with a dilute scheme can be used to evaluate the effective properties. For composites with a higher volume fraction of inclusions, aggregate model [12, 13] can be used, which utilizes the known properties of constituents. Mori-Tanaka [14] have proposed a theory by assuming that the average strain in the inclusion is

✉ A. Sharma
anujsharma_phd2k16@dtu.ac.in

[1] Department of Mechanical, Production, Industrial and Automobile Engineering, Delhi Technological University, Delhi 110042, India

Published online: 23 November 2020

related to the average strain in the matrix by a fourth-order tensor. The relation between the uniform strains in the inclusion in the matrix is provided by this fourth-order tensor. It was further extended by Benveniste [15] to investigate the stress and strain concentration tensors along with the overall elastic modulus of composites. This mean-field homogenization theory has been adopted by various researches [16–20] to determine the effective properties such as modulus, density and thermal expansion and is used in this work.

Significant increase of tensile strength, yield strength and hardness on the addition of alumina particles in the aluminium matrix is validated experimentally in archival literature [21–23]. Properties such as young's modulus, poisson's ratio and shear modulus are computationally derived successfully by various researchers. However, experimental validation of these effective properties with standard practice and technique is equally important. These properties can be experimentally determined by impact hammer test based on ASTM E1876 standard [24].

Recently, *Praveen* et al. [25] performed an impact hammer test based on ASTM E1876 to evaluate the dynamic elastic properties of composites at ambient temperatures. However, the comparison of multi-scale computation of the effective properties and the respective experimental results have not been explored in detail. Daramola et al. [26] compared the mean-field homogenization technique with the experiment and estimated an error of 4.2% between the predicted and experimental values. However, the microscopic parameters such as volume fraction were assumed to be same as the fraction of fibres added during manufacturing. For metal matrix composites fabricated by stir casting method, there is much difference between the added alumina particles and the particles mixed with the matrix. Therefore, a unified and novel approach is adopted in this paper, where the effective properties of the composite are derived computationally using parameters derived through an image analysis method. Moreover, these computationally driven values are compared with the experimental results as per ASTM E1876 standard, evaluated by the impact hammer test.

Material and Method

The composite shaft samples (800 mm length and 22 mm diameter) are fabricated using a stir casting process, as shown in Fig. 1. The aluminium 6061 metal matrix is mixed through stirring with a different weight percentage of alumina (0 wt.%, 6 wt.% and 12 wt.% of alumina) supplied by Loba Chemie Pvt. Ltd. In order to increase the wettability of the alumina particles in the aluminium matrix, the alumina particles were preheated, and magnesium was added during the stir casting process. Magnesium forms an interface of MgO and $MgAl_2O_4$ between the alumina and aluminium, resulting in better wettability and strong bonding [27]. As shown in Fig. 2, the bonding between the reinforcement and the aluminium matrix is strong for both dispersed and agglomerated particles. The Scanning Electron Microscopy (SEM) images are taken for these samples, and then these images are analyzed through an image analysis method. The average aspect ratio and average volume fraction are determined through image analysis. These parameters are used to predict the effective properties of the composite through Mori-Tanaka mean-field homogenization. These properties are then compared with the experimental response of the composite specimen.

Estimation of Effective Mechanical Properties through Mean-Field Homogenization by Using Parameters Extracted through Image Processing

Cubic samples of 5 mm length, 5 mm width and 5 mm height were processed by EDM wire cutting and Scanning Electron Microscopy (SEM) micrographs at the magnification of 500X

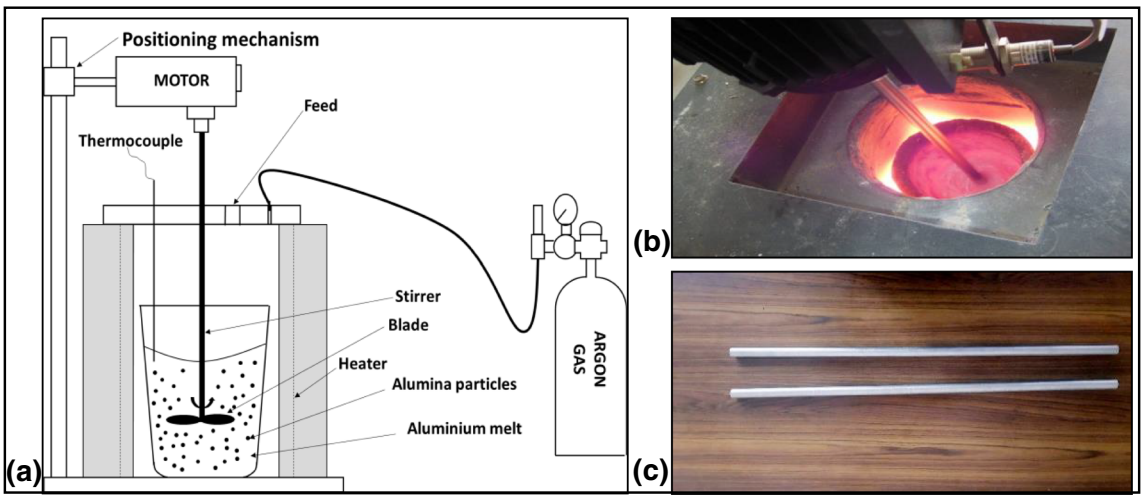


Fig. 1 (**a**) Stir casting set up (**b**) Vortex creation by a four-blade stirrer (**c**) Stir casted 800 x 22 mm shaft

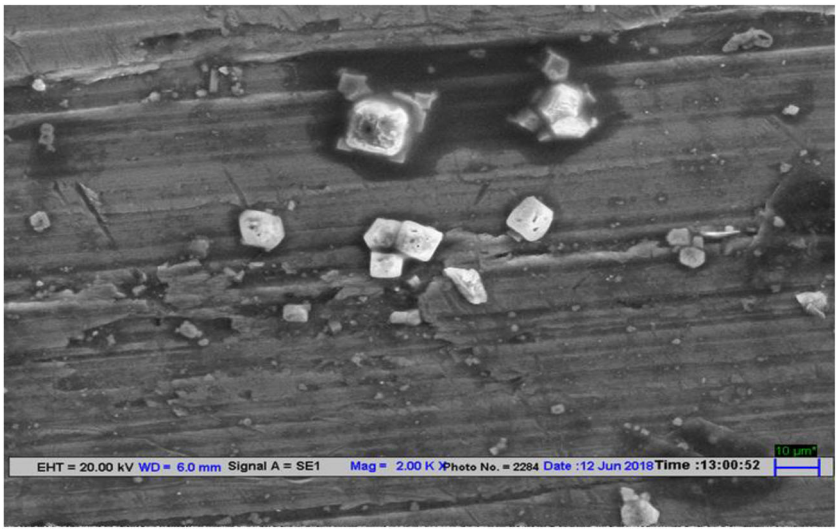


Fig. 2 2000X image showing strong bond between alumina particles and aluminium matrix

and 1000X were obtained using a Zeiss EVO 50 machine. Morphology of the micrographs depicts uniform distribution of reinforcements in the aluminium matrix. For each composition, three samples were scanned, and the fraction and size of alumina particles were evaluated by image processing and averaged for the computation of effective properties.

The open-source image processing software ImageJ was used to generate binary images of the composites with ellipsoid form, assumed for homogenization of properties. SEM micrographs were converted to binary images by using threshold of 61 foreground pixels and 255 background pixels. This transformed the different luminous pixels into white/black bifurcating with respect to the threshold values. At this value of threshold, particles were easily differentiated from the matrix material. Parameters such as the size of the reinforcement, aspect ratio and volume percentage for the composite were derived from the image analysis, as shown in Fig. 3. These parameters were incorporated in the mean-field homogenization of the metal matrix composite material where a Representative Volume Element (RVE) is assumed to be the building block of the macroscopic structure.

RVE is a statistical representation that connects microscopic structure and constituents with the macroscopic properties of composite materials. Therefore, the size of the RVE should

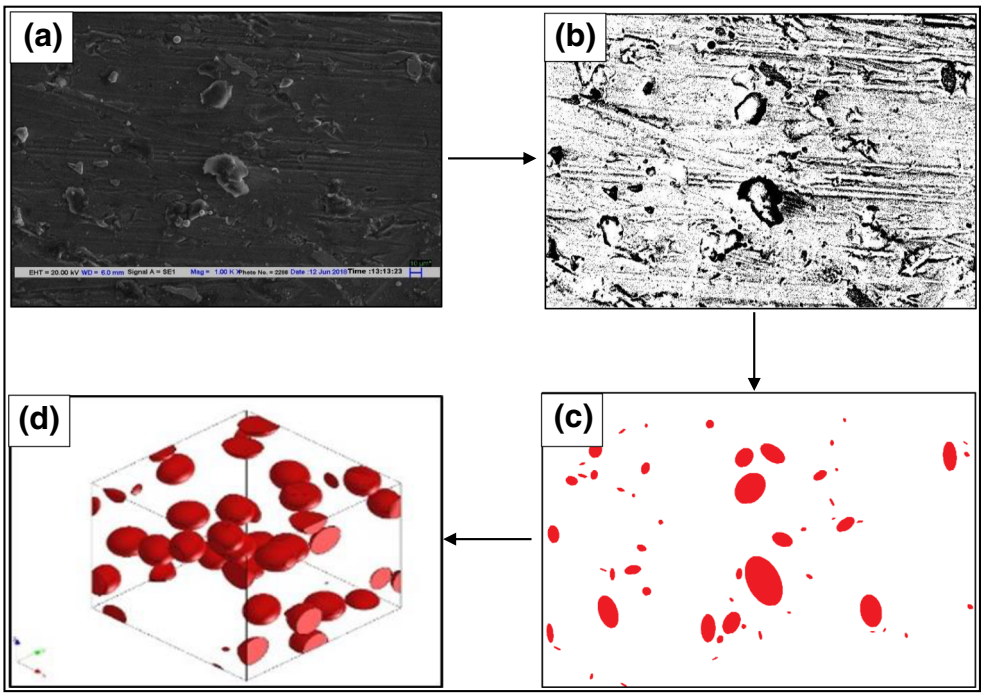


Fig. 3 Conversion of an SEM image into an RVE (**a**) SEM Image of a composite sample (**b**) Cropped SEM image converted into a binary image (**c**) Particle analysation of the binary image (**d**) RVE with mean aspect ratio, size and volume fraction

be such, that it is aptly smaller than the macroscopic structural dimensions and simultaneously large enough to contain an ample amount of information on the microstructure [28, 29].

Digimat is a composite modelling platform which has been used for the generation of RVEs with standard size and threshold. It consists of a Digimat-FE tool, which generates realistic RVEs with the finite element method by using the RVE generation algorithm. This algorithm is based on Random sequential adsorption (RSA) technique which randomizes the orientation of the particles [30]. Particles position themselves in random directions sequentially, and each particle blocks nearby locations from being occupied by another particle, therefore avoiding coagulation. This platform has been used for the development of RVEs in many types of researches. RVEs with randomly distributed spherical inclusions [31, 32], perfectly aligned fibres [33] or random fibre distributions by using a fibre randomization algorithm [34] are generated by using this technique. These RVEs can be exported as standard CAD formats. The effective properties of these RVEs were calculated and have been used for further analysis.

For mean-field homogenization, Mori-Tanaka method [14] has been adopted to determine the effective Young's modulus of the composite material. Therefore, it assumed that the metal matrix composite is comprised of N number of phases. Matrix phase is denoted by subscript m, and there are remaining N-1 inclusion phases in the control volume. The volume fractions of the matrix and the i_{th} inclusion are denoted by ϑ_m and ϑ_i. Similarly, stiffnesses of the matrix and the i_{th} inclusion are denoted by K_m and K_i. These stiffness quantities are generally represented as elasticity tensors of fourth-order which possess specific symmetrical properties. The overall elastic stiffness of the composite [35] is defined in Eq. (1).

$$\bar{\bar{K}}_{av} = \left(\vartheta_m * K_m + \sum_{i=1}^{N-1} \vartheta_i \langle K_i * D_i^{dil} \rangle\right) + \left(\vartheta_m * \boldsymbol{I} + \sum_{i=1}^{N-1} \vartheta_i \langle D_i^{dil} \rangle\right)^{-1} \tag{1}$$

Where $\boldsymbol{I}$ is the identity tensor and the dilute strain concentration factor D_i^{dil} is generally expressed in Eq. (2).

$$D_i^{dil} = \left[\boldsymbol{I} + S_i\, K_m^{-1}(K_i - K_m)\right]^{-1} \tag{2}$$

Here S_i is the Eshelby tensor [6] for the spheroidal inclusion in the infinite matrix.

Since the alumina particles are randomly oriented, orientation averaging has to be taken into account. Therefore, the terms in the angle parentheses ⟨⟩ denote the averaged value of the parameters over all possible orientations. This orientationally averaged tensor is a fourth-order tensor in the 3D space [35] and can be expressed as in Eq. (3).

$$\langle T_{klmn} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} T_{klmn}(\alpha, \beta) sin\beta \, d\alpha \, d\beta \tag{3}$$

where α, β are the generalized angular orientations. For the randomization, the local coordinates should be transformed to the global coordinates. For overall random transformation, the matrix may be expressed as sown in Eq. (4).

$$a_{kl} = \begin{bmatrix} \cos\alpha & \cos\beta \sin\alpha & \sin\beta \sin\alpha \\ -\sin\alpha & \cos\beta \cos\alpha & \sin\beta \cos\alpha \\ 0 & -\sin\beta & \cos\beta \end{bmatrix} \tag{4}$$

Moreover, the transformation may be written, as shown in Eq. (5).

$$T_{klmn}(\alpha, \beta) = a_{ko} a_{lp} a_{mq} a_{nr}\, T'_{opqr} \tag{5}$$

The effective Young's Modulus(E_{efv}) [35] can be now calculated by the general equation (Eq. (6)) with shear modulus and bulk modulus.

$$E_{efv} = E_m * \frac{\kappa_{efv} * \mu_{efv}(3\kappa_m + \mu_m)}{3k_{efv} * \kappa_m + \mu_{efv} * \mu_m} \tag{6}$$

Effective Shear modulus can be derived using Eq. (7).

$$\frac{\kappa_{efv}}{\kappa_m} = \frac{1}{1 + z\, p} \tag{7}$$

Effective Bulk modulus can be derived using Eq. (8).

$$\frac{\mu_{efv}}{\mu_m} = \frac{1}{1 + z\, q} \tag{8}$$

Where E_{efv} is the effective Young's modulus, E_m is Young's modulus of the matrix, κ_{efv} and μ_{efv} are the effective bulk and shear moduli of the composite, z is the volume concentration/fraction, κ_m and μ_m are the bulk and shear modulus of the matrix. Also, p and q are the explicit expressions, and these expressions and details of the expression can be seen in Tandon and Weng [36] for spheroidal inclusions for detailed reference.

Also, effective density can be easily calculated through the rule of the mixture, as shown in Eq. (9).

$$\rho_{efv} = \rho_m \vartheta_m + \rho_i \vartheta_i \tag{9}$$

Where ρ_{efv} is the effective density, ρ_m and ϑ_m are the density and volume fraction of matrix and ρ_i and ϑ_i are the density and volume fraction of inclusions.

Experimental Evaluation of Elastic Properties

For the experimental validation of the estimated/predicted properties of the composites, ASTM E1876 impact hammer vibration analysis was employed. The analysis is performed on a standard experiment setup which conforms the prerequisites for the ASTM E1876 standard, as shown in Fig. 4.

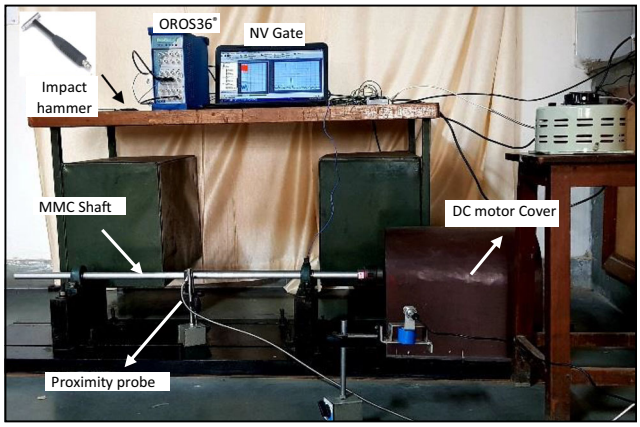


Fig. 4 Experimental set up for impact hammer test

The inductive proximity sensor (LJ12A3–4-Z /BY) having a sensitivity of 1000 V/m and a detection range of 4000 μm is used. The test specimens are energized through a force transducer with a plastic tip. Impact Hammer (Model: - PCB-086C03) with sensitivity 2250 μ(V)/(N) and range 4440 N is used for transient excitation. The response from the transducers is collected and analyzed by using a 32 channel, vibration analyzer OROS36®. The channels of these analyzers are handled in real-time: FFT, 1/3rd Octave. The fundamental natural frequencies of the specimen are observed from the peaks of the frequency response function obtained by using NV Gate® (version 10.1.1) interface.

The prime objective of the experimentation was to extract the values of natural frequencies for the first two bending (flexural) modes of the shaft. A range of 0–600 Hz was targeted as per the dimensions for experimental analysis to confirm that primarily these modes are excited. For excitation, the plastic tip of suitable hardness was used such that the input spectrum doesn't prominently excite the frequencies beyond the working range. The experimental samples were selected carefully such that the impact force input spectrum excites the required range of frequency at amplitude high enough to avoid noise interceptions. The sampling rate of 2.04 kS/s (kilo samples per second) was chosen, which was much higher than the Nyquist frequency according to the range excited by the input spectrum and the range of observation. This high rate of sampling prevented the problem of leakage and aliasing, and no anti-aliasing filters were required. The motive of the experimentation was also to observe the damping of the material relative to each other. The windowing technique would have altered the peak width and thus, the estimation of damping. Therefore, the time range was taken in such a way that windowing was not required for sample selection. The high resolution of the sampling ensured that the natural frequency is confidently determined by Gaussian peak fitting method. The position of the sensor has to be chosen in such a way that it should detect the first and second mode excitation. At the centre of the shaft, the first mode was highly excited, but the second mode was not excited (Since the antinode of the second mode is at the centre for the simply supported system). Therefore, a non-contact inductive proximity transducer was roved, and a suitable position was chosen where the excitation of first and second peaks was obtained successfully (At a distance of 50 mm from the centre as shown in Fig. 5. At this position, the amplitude of first bending natural frequency peak is the highest followed by second bending natural frequency peak.

The Young's modulus expression for the known fundamental flexural natural frequency is shown in Eq. (10).

$$E = 1.6067\left(\frac{l^3}{d^4}\right)\left(mf_f^2\right)T_1 \tag{10}$$

where E = Young's modulus, l = Length of the specimen, d = diameter of the shaft, m = mass of the shaft, f_f = fundamental

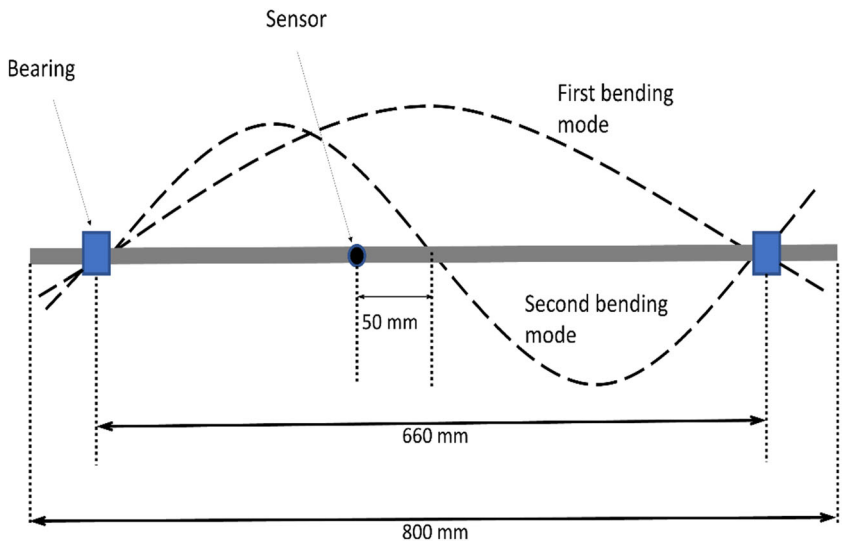


Fig. 5 First and second bending mode with sensor positioning

Table 1 Mean volume fraction, aspect ratio and particle size for various SEM images

Shaft Material		Mean volume fraction			Mean aspect ratio			Mean particle size (μm)		
		S1	S2	S3	S1	S2	S3	S1	S2	S3
Al 6061/6 wt.% Al_2O_3	Max	4.0	4.8	4.7	3.1	2.9	3.1	8.7	8.9	8.8
	Min	2.9	3.4	3.1	2.6	2.6	2.7	7.3	7.2	7.1
	Mean	3.4	4.1	3.8	2.9	2.7	3.0	8.0	8.1	8.1
Al 6061/12 wt.% Al_2O_3	Max	8.3	9.4	9.9	3.4	3.2	3.3	8.9	8.8	9.0
	Min	6.9	7.4	7.9	2.8	2.7	2.9	7.3	7.5	7.4
	Mean	7.3	8.1	9.1	3.2	3.0	3.1	8.3	8.2	8.4

flexural frequency of the shaft and T_1 = correction factor for the fundamental flexural mode to account for finite diameter of the rod, and Poisson's ratio.

Since the aspect ratio (l/d) is higher than 20, the simplified form for T_1 can be used as shown in Eq. (11).

$$T_1 = \left[1 + 4.939\left(\frac{d}{l}\right)^2\right] \tag{11}$$

The expression for Dynamic shear modulus for known fundamental torsion resonant frequency of a cylindrical is shown in Eq. (12).

$$G = 16mf_t^2\left(\frac{l}{\pi d^2}\right) \tag{12}$$

where G = Dynamic shear modulus, l = Length of the specimen, d = diameter of the shaft, m = mass of the shaft, f_t = fundamental torsional frequency of the shaft.

Correspondingly, Poisson's Ratio can be determined using Eq. (13).

$$\mu = \left(\frac{E}{2G}\right) - 1 \tag{13}$$

Where μ is the Poisson's ratio.

Results and Discussion

Since there might be inhomogeneous zones within the sample which lead to interference of the results, measures are taken to avoid these uncertainties. Three samples were taken for each material for SEM analysis. The SEM images were randomly taken for different positions within the sample and all six surfaces of the cuboid sample with 1000x magnifications. For each sample (S1, S2 and S3), maximum, minimum, and mean values of volume fraction, aspect ratio and particle size are taken for various SEM micrographs as shown in Table 1.

The aspect ratio and particle size doesn't vary much for different samples and materials. The mean volume fraction increased for the Al 6061/12 wt.% Al_2O_3 as compared to the Al 6061/6 wt.% Al_2O_3. The variation of the microstructural parameters within the sample were not resulting in interference with the microstructural parameters of different composition. For example, the mean volume fraction ranges from 2.9 to 4.8 for Al 6061/6 wt.% Al_2O_3 and 6.9 to 9 for Al 6061/12 wt.% Al_2O_3. The data is therefore categorical, and we can classify the results for different composition of materials respectively.

The modulus of elasticity is highest for the largest volume fraction, highest aspect ratio and smallest mean particle size and vice versa. Therefore, the overall highest young's modulus can be predicted by taking mean volume fraction as 4.8, aspect ratio as 3.1 and mean particle size as 7.1 μm and vice versa for Al 6061/6 wt.% Al_2O_3. Also, overall highest young's modulus can be predicted by taking mean volume fraction as 9.7, aspect ratio as 3.4 and mean particle size as 9.0 μm and vice versa for Al 6061/12 wt.% Al_2O_3. The consolidated results for the estimated microscopic properties through image analysis and mean-field homogenization are shown in Table 2.

On the addition of 6 wt.% of alumina, the overall mean volume fraction of 3.8% is observed by image analysis of

Table 2 Image analysis overall mean results

Material	Mean volume fraction	Mean aspect ratio	Mean particle size (μm)
Al 6061	–	–	–
6 wt.% Al_2O_3/Al 6061	3.8	2.9	8.1
12 wt.% Al_2O_3/Al 6061	8.166	3.1	8.3

Table 3 Mean-field homogenization results

Engineering constants	Al 6061	6 wt.% Al_2O_3/Al 6061			12 wt.% Al_2O_3/Al 6061		
		Mean	Max	Min	Mean	Max	Min
In-Plane Young's modulus (GPa)	70	73.73	74.84	72.9	78.66	80.1	77.4
Out -Plane Young's modulus (GPa)	70	73.61	74.71	72.78	78.32	79.75	77.06
In-Plane Poisson's ratio	0.328	0.326	0.325	0.326	0.322	0.321	0.323
Transverse Poisson's ratio	0.328	0.326	0.325	0.326	0.322	0.321	0.323
In-Plane shear modulus (GPa)	26.35	27.8	28.21	27.48	29.75	30.29	29.27
Transverse shear modulus (GPa)	26.35	27.76	28.17	27.44	29.2	29.73	28.73
Global Density (Kg/m^3)	2710	2748	2772	2723	2798	2838	2753

SEM images for different samples. The observed volume fraction of alumina in fabricated samples is low as compared to the volume fraction of alumina added during stir casting. The limitation of stir casting method that all alumina particles are not mixed with the matrix is the reason for this deviation. Although stir casting is flexible enough and large-sized samples used for analysis can be easily manufactured through this method. Similarly, for the addition of 12 wt.% of alumina particles, the mean volume fraction of 8.166% is observed. The mean aspect ratio and the average particle size are also in the conformance with the alumina particles manufacturer specification.

The effective Young's Modulus, Poisson's ratio and the shear modulus of in-plane and transverse plane for a particular RVE is estimated through mean-field homogenization as shown in Table 3.

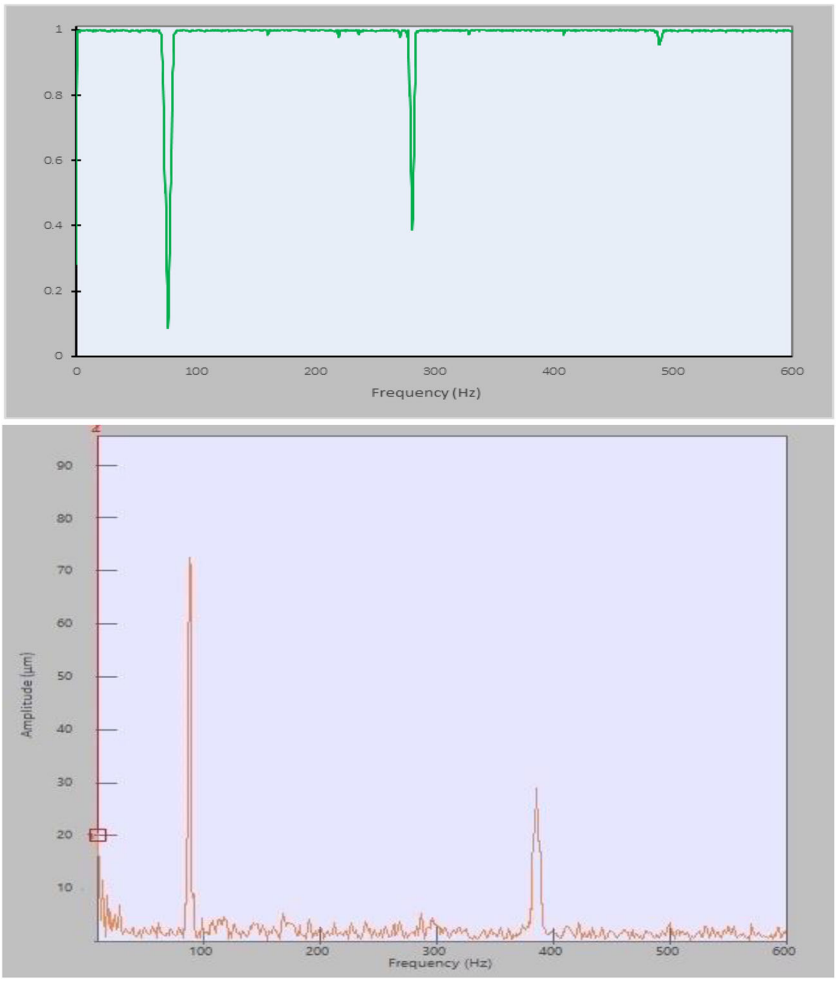


Fig. 6 FRF and coherence for Al6061

The estimated mean In-plane Young's modulus is increased by 5.32%, on the addition of 6 wt.% of alumina due to enhanced dislocation density and precipitation hardening. Since the orientation of the particles is randomized, the value of out of plane Young's modulus is approximately the same as values of In-plane Young's modulus. The global density of the composite is also estimated to be increased by the addition of the particles since the density of the alumina particles is more than the aluminium matrix. The mean In-Plane shear modulus is also increased by 5.5% with transverse shear modulus value approximately similar to the In-plane shear modulus. Moreover, the effective In-plane Young's Modulus is estimated to be increased by 12.37% on addition of 12 wt.% of alumina. The highest value of density and shear modulus is observed for the composite with 12 wt.% of alumina reinforcements.

As shown in Table 3, there is no interference in respective values of Young's modulus, Poisson's ratio and shear modulus for different materials. For example, the minimum value of In-plane young's modulus of 12 wt.% Al_2O_3/Al 6061 is 77.4 GPa, whereas the maximum value of In-plane young's modulus of 6 wt.% Al_2O_3/Al 6061 is 74.84 GPa. So, this data can be classified by the composition of the material.

The strengthening effect of reinforcement is not directionally dependent as in the case of fibre reinforced composites due to the random orientation of particles. By adding 12 wt.% of alumina, In-Plane Young's modulus is increased by 12.37%, but the global density is increased by 3.2%. It implies that the addition of reinforcement improves the specific modulus of the material and therefore improves the modal response of the material. Therefore, impact hammer test is performed to obtain the properties of composites through modal response.

The frequency response function and the coherence for Al6061, Al6061/6 wt.% Al_2O_3 and Al6061/ 12 wt.% Al_2O_3 composite in the impact hammer test is shown in Figs. 6, 7 and 8. The input and output spectrums were coherent for all shafts with the resonance value dropped at anti-nodes. The high coherence between the input and output spectrum, except at the anti-nodes confirms the replicability of the experiments. The high resolution of the frequency spectrum was obtained because of the high sampling rate. The amplitude and the natural frequency were calculated by determining the peak using the Gaussian peak fitting method, and the damping ratio is obtained by the half-power bandwidth method.

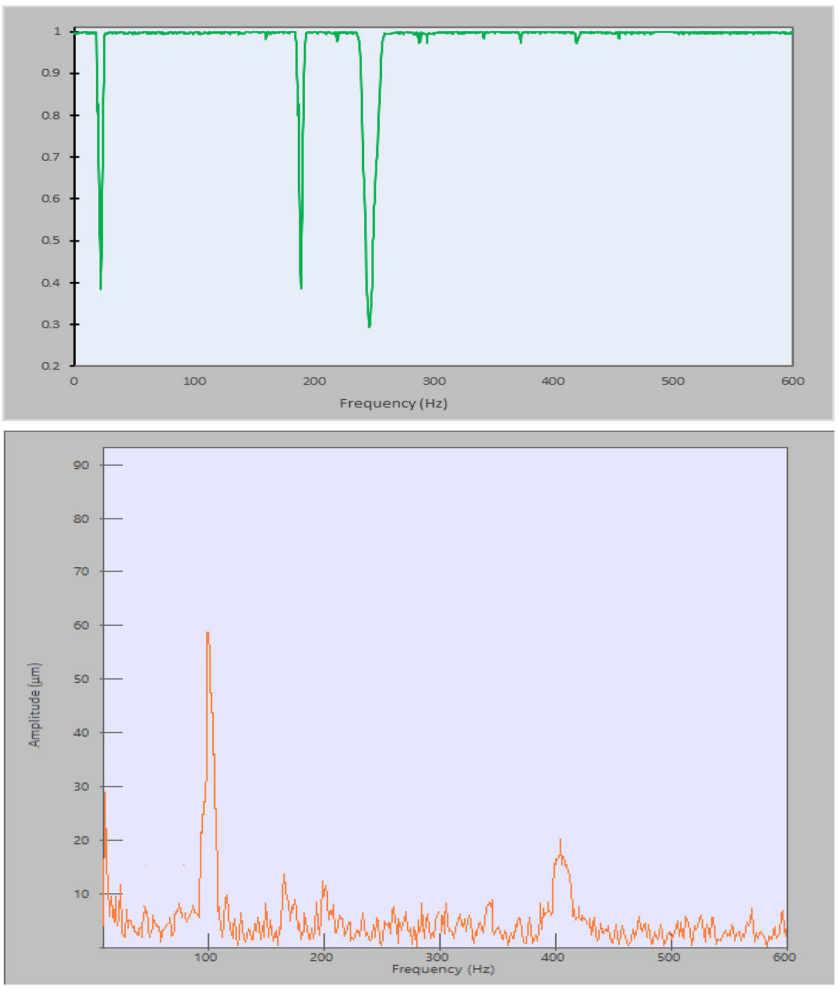


Fig. 7 FRF and coherence for Al6061/ 6 wt. % Al_2O_3

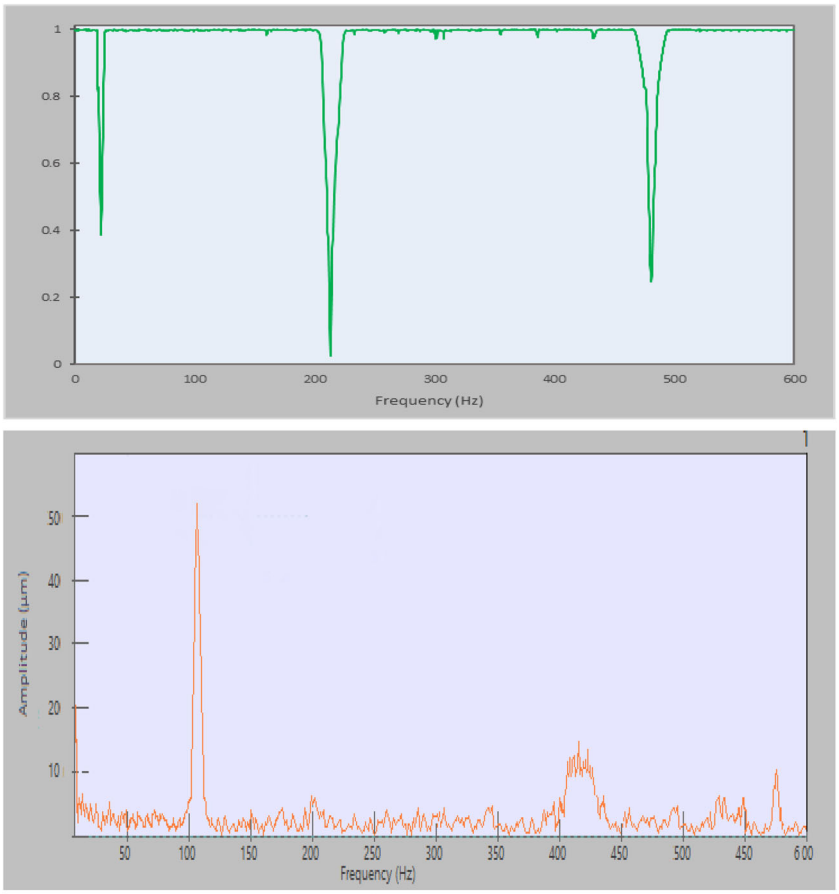


Fig. 8 FRF and coherence for Al6061/ 12 wt. % Al_2O_3

For each shaft with different composition, thirty experiments were taken for the determination of natural frequencies, damping ratio and amplitudes. The mean, standard deviation and range of the natural frequencies for different sets of experimentations are shown in Table 4.

The low standard deviation was obtained for the values of natural frequencies for both first and second mode. Also, the range of natural frequency values for different composition doesn't interfere with each other. The natural frequency increases significantly with the increase of reinforcement particles in the shafts. This is due to the reason that the addition of alumina particles in the Al6061 matrix leads to a substantial improvement in the bending stiffness as compared to the material density. The mean, standard deviation and range of the damping for different sets of experimentations are shown in Table 5.

The higher value of damping was observed for the specimen with higher wt.% of alumina particles in the aluminium matrix. The damping was significantly increased for the second mode of vibration. It suggests that the damping increases with the increase of alumina particles in the aluminium matrix. This is due to the grain refinement and formation of secondary phases which increases elastic strain energy dissipation in the materials.

The mean, standard deviation and range of the amplitude of vibrations for different sets of experimentations are shown in Table 6.

Table 4 Experimental natural frequency mean, standard deviation and range

Shaft Material	Natural frequency (first mode)			Natural frequency (second mode)		
	Mean	Standard deviation	Range	Mean	Standard deviation	Range
Al6061	98.62	0.65	97.6–99.1	388.93	1.92	386.1–391.6
Al 6061/6 wt.% Al_2O_3	101.65	0.72	100.1–102.8	408.43	2.42	406.3–411.7
Al 6061/12 wt.% Al_2O_3	108.91	1.01	107.1–110.7	420.21	3.02	416.2–425.1

Table 5 Experimental damping mean, standard deviation and range

Shaft Material	Damping (first mode)			Damping (second mode)		
	Mean	Standard deviation	Range	Mean	Standard deviation	Range
Al6061	0.0097	0.0001	0.0095–0.0099	0.021	0.0003	0.0207–0.026
Al 6061/6 wt.% Al_2O_3	0.0123	0.0004	0.0119–0.0129	0.0233	0.0005	0.0225–0.0242
Al 6061/12 wt.% Al_2O_3	0.0146	0.0006	0.0137–0.155	0.026	0.0008	0.0252–0.0262

Table 6 Experimental amplitude mean, standard deviation and range

Shaft Material	Amplitude (first mode) (μm)			Amplitude (second mode) (μm)		
	Mean	Standard deviation	Range	Mean	Standard deviation	Range
Al6061	69.1	2.11	67.1–70.9	30.21	2.01	31.2–27.1
Al 6061/6 wt.% Al_2O_3	62.7	2.03	60.7–64.2	21.62	2.42	22.5–19.1
Al 6061/12 wt.% Al_2O_3	52.1	2.81	50.2–54.8	14.29	2.65	12.31–17.01

The amplitude reduced significantly with the increase in alumina particles for both the modes. The range and the standard deviation are well within the limit that is required for the data to be categorical about the composition of the material. This is due to the increase in bending stiffness of the shafts because of the reinforcing effect of the alumina particles in the matrix.

These values are utilized to obtain the value of In-plane Young's modulus, shear modulus and Poisson's ratio experimentally, as per ASTM E1876 standard. The comparison of predicted values obtained through mean-field homogenization and the experimental values obtained is shown in Table 7.

The experimental value of In-plane Young's modulus of Al 6061 is observed to be 2.88% lower than that of initial prediction. This lower value can be attributed to the defects induced during the casting process of aluminium alloy. However, these defects might have prevailed for the case of 6 wt.% Al_2O_3/Al 6061 also, but the experimental In-plane Young's modulus has increased by 1.19% as compared to the predicted value. Similarly, experimental In-plane Young's modulus has increased by 3.75% as compared to the predicted value for the case of 12 wt.% Al_2O_3/Al 6061. From literature [26], it is observed that the addition of 6 wt.% micro-sized kaolinite particle in epoxy polymer matrix composite increased the effective elastic modulus from 3100 MPa to 3107.5 MPa. The mean-field homogenization method predicted the value 4.2% higher than the experimental value. Whereas in the present study, the mean-field homogenization predicted the value 1.19% lower than the experimental values. The reason for predicting a lower value by the homogenization process may be that the actual alumina particle volume fraction is higher than the volume fraction estimated by image analysis.

The range of the experimental and the predicted values for the In-plane Young's modulus is shown in Fig. 9. For Al 6061, the predicted value is the standard value for Young's modulus, and the experiment values are of the cast material. Therefore, the values are lower due to the introduction of some porosity. For, Al 6061/ 6 wt.% Al_2O_3 the experimental range intersects with the predicted value range. For, Al 6061/ 12 wt.% Al_2O_3 the experimental range exceeds the predicted

Table 7 Comparison of Predicted and experimental value

Engineering constants	Al 6061		6 wt.% Al_2O_3/Al 6061		12 wt.% Al_2O_3/Al 6061	
	Predicted	Experimental	Predicted	Experimental	Predicted	Experimental
In-Plane Young's modulus	70 GPa	67.98 GPa	73.73 GPa	74.61 GPa	78.66 GPa	81.61 GPa
In-Plane shear modulus	26.35 GPa	25.55 GPa	27.82 GPa	29.01 GPa	29.75 GPa	30.91 GPa
In-Plane Poisson's ratio	0.328	0.327	0.326	0.323	0.322	0.320

Fig. 9 Comparison of range of predicted and experimental value of In-plane Young's modulus

value range at a few points. Therefore, it can be concluded that the upper bounds of mean-field solutions should be used for the higher reinforcement compositions.

The similar kind of trend is observed for In-plane shear modulus, and the mean-field homogenization analysis does not deviate significantly from the experimental results.

Conclusion

The composite specimens have been successfully fabricated by stir casting method as per ASTM E1876 standards. The microstructure of the composites was taken through SEM and particles average size; aspect ratio and fraction were obtained through image analysis method. The effective properties were then evaluated by mean-field homogenization method by using these parameters. These properties were then evaluated experimentally as per ASTM E1876 standards. Following conclusions can be driven from the analysis in this paper:

- The volume fraction of the alumina particles observed by the image analysis method was significantly less as compared to the added volume fraction during fabrication. It suggests that the volume fraction added initially during fabrication should not be used directly for predicting the effective properties.
- The addition of alumina particles results in the strengthening of composite specimens due to enhanced dislocation density and precipitation hardening.
- The difference between the In-Plane and transverse effective properties is negligible since the strengthening effect of reinforcement is not directionally dependent due to the random orientation of particles.
- The addition of alumina particles in the Al6061 matrix leads to a substantial improvement in the bending stiffness as compared to the material density. Due to this reason, the natural frequency for various modes of vibration increased on the addition of particles.
- Damping ratio also increased on the addition of alumina particles due to the grain refinement and formation of secondary phases.
- The estimated error for the case of 6 wt.% Al_2O_3/Al 6061 is 1.19%, whereas the estimated error for the case of 6 wt.% Al_2O_3/ Al6061 is 3.75% for the mean values.
- The range of mean-field values lags behind the experimental values as the material of high reinforcement composition is used. It can be concluded that the upper bound of mean-field homogenization can be used for the material of high reinforcement composition and can be the future scope of work. However, mean-field homogenization predicts a low higher than experimental value while assuming that all alumina particles added for process are mixed into the matrix.
- Also, samples were taken from the end of the shaft for SEM analysis; this might be the low/high composition area of the shaft during casting. This variation can be addressed in future work.
- The variation of the experimental and the predicted values was such that categorization of the values was possible in reference to the composition of the shaft.

Author's Contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Mr. Anuj Sharma, Prof. Vikas Rastogi and Prof. Atul kumar agrawal The first draft of the manuscript was written by Mr. Anuj Sharma and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they do not have any conflict of interest.

References

1. Chawla N, Shen YL (2001) Mechanical behavior of particle reinforced metal matrix composites. Adv Eng Mater 3(6):357–370. https://doi.org/10.1002/1527-2648(200106)3:6<357::AID-ADEM357>3.0.CO;2-I
2. Rawal S (2001) Metal-matrix composites for space applications. Jom 53(4):14–17. https://doi.org/10.1007/s11837-001-0139-z
3. Sarmasti AB, Yazdanirad M, Nouri Khezrabad M, Karbalaie M (2011) Effect of alumina particle size and thermal condition of casting on microstructure and mechanical properties of stir cast Al-Al2O 3 composites. Mater Sci Technol 27(11):1653–1656. https://doi.org/10.1179/1743284710Y.0000000020
4. Koczak MJ, Khatri SC, Allison JE, Bader MG (1993) Metal-Matrix Composites for Ground Vehicle, Aerospace, and Industrial Applications. Fundamentals of Metal-Matrix Composites, 297–326. https://doi.org/10.1016/b978-0-08-052371-2.50020-1
5. Singh S, Singh I, Dvivedi A Design and development of novel cost-effective casting route for the production of metal matrix composites (MMCs) (2017). Int J Cast Metals Res 6(30):356–364. https://doi.org/10.1080/13640461.2017.1323605

6. Chaturvedi SK, Tzeng GY (1991) Micromechanical modeling of material damping in discontinuous fiber three-phase polymer composites. Compos Eng 1(1):49–60. https://doi.org/10.1016/0961-9526(91)90025-N
7. Guden M, Hall IW (1998) Dynamic properties of metal matrix composites: a comparative study. Mater Sci Eng A 242(1–2): 141–152. https://doi.org/10.1016/s0921-5093(97)00516-9
8. Lesieutre GA, Yarlagadda S, Yoshikawa S, Kurtz SK, Xu QC (1993) Damped structural composite material using resistively-shunted piezoelectric ceramic short fibers. Collect Tech Papers - AIAA/ASME Structures, Structural Dynamics Mater Conf (pt 6): 3238-3243
9. Vaidya UK, Dadzie P-E, Haque A, Mahfuz H, Jeelani S (1997) Nondestructive evaluation and characterization of microfibers modified textile carbon-phenolic and carbon-carbon composites. J Reinf Plast Compos 16(11):968–1001. https://doi.org/10.1177/073168449701601101
10. Mura T (2013) Micromechanics of defects in solids. Mechanics of Elastic and Inelastic Solids. Springer Science & Business Media. https://doi.org/10.1007/978-94-009-3489-4
11. Eshelby JD (1957) The determination of the elastic field of an ellipsoidal inclusion, and related problems. Proc Royal Soc A: Math Phys Eng Sci 241(1226):376–396. https://doi.org/10.1098/rspa.1957.0133
12. Christensen RM, Waals FM (1972) Effective stiffness of randomly oriented fibre composites. J Compos Mater 6(3):518–535. https://doi.org/10.1177/002199837200600307
13. Halpin JC, Jerine K, Whitney JM (1971) The laminate analogy for 2 and 3 dimensional composite materials. J Compos Mater 5(1):36–49. https://doi.org/10.1177/002199837100500104
14. Mori T, Tanaka K (1973) Average stress in matrix and average elastic energy of materials with misfitting inclusions. Acta Metall 21(5):571–574. https://doi.org/10.1016/0001-6160(73)90064-3
15. Benveniste Y (1987) A new approach to the application of Mori-Tanaka's theory in composite materials. Mech Mater 6(2):147–157. https://doi.org/10.1016/0167-6636(87)90005-6
16. Abaimov SG, Trofimov A, Sergeichev IV, Akhatov IS (2019) Multi-step homogenization in the Mori-Tanaka-Benveniste theory. Compos Struct. https://doi.org/10.1016/j.compstruct.2019.03.073
17. Imani SM, Goudarzi AM, Rabiee SM, Dardel M (2018) The modified Mori-Tanaka scheme for the prediction of the effective elastic properties of highly porous ceramics. Ceram Inter 44(14):16489–16497. https://doi.org/10.1016/j.ceramint.2018.06.066
18. Kamarian S, Shakeri M, Yas MH (2016) Natural frequency analysis and optimal design of CNT/fiber/polymer hybrid composite plates using Mori-tanaka approach, GDQ technique, and firefly algorithm. Polym Compos 39(5):1433–1446. https://doi.org/10.1002/pc.24083
19. Tran VP, Brisard S, Guilleminot J, Sab K (2018) Mori–Tanaka estimates of the effective elastic properties of stress-gradient composites. Int J Solids Struct 146:55–68. https://doi.org/10.1016/j.ijsolstr.2018.03.020
20. Sobhaniaragh B, Batra RC, Mansur WJ, Peters FC (2017) Thermal response of ceramic matrix nanocomposite cylindrical shells using Eshelby-Mori-Tanaka homogenization scheme. Compos Part B 118:41–53. https://doi.org/10.1016/j.compositesb.2017.02.032
21. Yılmaz O, Buytoz S (2001) Abrasive wear of Al_2O_3-reinforced Aluminium-based MMCs. Compos Sci Technol 61:2381–2392
22. Kok M (2005) Production and mechanical properties of Al_2O_3 particle-reinforced 2024 Aluminium alloy composites. J Mater Process Technol 161:381–387
23. Suthar J, Patel KM (2018) Processing issues, machining, and applications of aluminium metal matrix composites. Mater Manuf Process 33(5):499–527. https://doi.org/10.1080/10426914.2017.1401713
24. Standard Test Method for Dynamic Young's Modulus, Shear Modulus, and Poisson's Ratio by Impulse Excitation of Vibration. ASTM E1876, ASTM International, United States, 1–17
25. Praveen AP, Rajamohan V, Mathew AT (2019) Material and mechanical characterization of multi-functional carbon nanotube reinforced hybrid composite materials. Exp Tech 43(3):301–314. https://doi.org/10.1007/s40799-019-00316-0
26. Daramola OO, Olajide JL, Adediran AA, Adewuyi BO, Ayodele TT, Desai DA, Sadiku ER (2020) Multiscale analysis and experimental validation of the effective elastic modulus of epoxy-dioctahedral phyllosilicate clay composite. Heliyon 6(6):e04008. https://doi.org/10.1016/j.heliyon.2020.e04008
27. Sangghaleh A, Halali M (2009) Effect of magnesium addition on the wetting of alumina by aluminium. Appl Surf Sci 255:8202–8206. https://doi.org/10.1016/j.apsusc.2009.05.044
28. Bouhfid N, Raji M, Boujmal R, Essabir H, Bensalah M-O, Bouhfid R, Qaiss AeK. (2019) Numerical modeling of hybrid composite materials. Modelling of Damage Processes in Biocomposites, Fibre-Reinforced Composites and Hybrid Composites, 57–101. https://doi.org/10.1016/b978-0-08-102289-4.00005-9
29. Gitman IM, Askes H, Sluys LJ, Valls OL (2004) The concept of representative volume for elastic, hardening and softening materials. International Summer School - IPME conference. XXXII International Summer School - IPME Conference: Advanced Problems in Mechanics. p 1-5
30. Babu KP, Mohite PM, Upadhyay CS (2018) Development of an RVE and its stiffness predictions based on mathematical homogenization theory for short fibre composites. Int J Solids Struct 130-131:80–104. https://doi.org/10.1016/j.ijsolstr.2017.10.011
31. Klusemann B, Svendsen B (2010) Homogenization methods for multi-phase elastic composites: comparisons and benchmarks. Tech Mech 30(4):374–386
32. Yi I, Wiedmaier J, Schmauder S (2015) Modeling and simulation of mechanical properties of nano particle modified polyamide 6. J Mater Sci Chem Eng 3(1):80–87
33. Ogierman W, Kokot G (2016) A study on fiber orientation influence on the mechanical response of a short fiber composite structure. Acta Mech 227(1):173–183
34. Soni G, Singh R, Mitra M, Falzon BG (2014) Modelling matrix damage and fibre matrix interfacial decohesion in composite laminates via a multi-fibre multilayer representative volume element (M2RVE). Int J Solids Struct 51(2):449–461
35. Chang CY, Chang MY, Huang JH (2004) Vibration analysis of rotating composite shafts containing randomly oriented reinforcements. Compos Struct 63(1):21–32
36. Tandon GP, Weng GJ (1986) Average stress in the matrix and effective moduli of randomly oriented composites. Compos Sci Technol 27(2):111–132. https://doi.org/10.1016/0266-3538(86)90067-9

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/344843520

Feature Selection Based Twin-Support Vector Machine for the Diagnosis of Parkinson's Disease

Conference Paper · December 2020

CITATIONS
0

READS
14

4 authors, including:

Surendrabikram Thapa
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS

SEE PROFILE

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

Early Detection of Alzheimer's Disease View project

All content following this page was uploaded by Surendrabikram Thapa on 23 October 2020.

The user has requested enhancement of the downloaded file.

Feature Selection Based Twin-Support Vector Machine for the Diagnosis of Parkinson's Disease

Surendrabikram Thapa[1], Surabhi Adhikari[1], Awishkar Ghimire[1,#] and Anshuman Aditya[1]
[1]Department of Computer Science and Engineering, Delhi Technological University, India
[#]awishkar.ghimire@gmail.com

***Abstract*—With growing number of ageing population, Parkinson's disease has become a serious problem to huge fraction of people above 60. The disease severely affects the motor system and can lead to death of the patients. There is no cure available for the disease. The symptoms on motor system is seen very late which leads into difficulty in management of the disease. There is no cure for the disease which makes it more difficult when the disease is diagnosed later. For early stages of Parkinson's disease, there are some medications to improve the symptoms. There are certain symptoms like slurred speech, problems in utterances, etc. which are seen earlier. These symptoms can be leveraged to diagnose the disease in its earlier stages. Recently, computers and machine learning algorithms have been widely used in diagnosis of various diseases. Speech attributes can be analyzed using machine learning algorithms to build predictive models for detection of Parkinson's disease. In this paper, twin-support vector machine (TSVM) based on feature selection technique has been discussed along with other ML techniques for the early diagnosis of Parkinson's disease. Feature selection based TSVM has shown promising results in classification of Parkinson's disease patients from healthy subjects.**

Keywords—Parkinson's disease, feature selection, twin support vector machines, speech analysis

I. INTRODUCTION

Parkinson's Disease (PD) was discussed and described for the first time by James Parkinson in 1817 [1]. PD is a progressive central nervous system disorder that mostly affects the motor system [2]. It affects the patient's ability with movement. The exact cause of Parkinson's disease is yet to be known but it is mostly caused by a loss of nerve cells in the part of the brain called the substantia nigra [3]. It is a progressive disease that gradually starts with stiffness in muscles, barely noticeable tremors in hands, and eventually affecting the whole movement [4]. This eventually leads to loss of automatic unconscious movements like smiling, blinking of the yes or swinging of arms while walking. It is also characterized by mild to severe depression [5].

The effects of PD in the motor system of the patients are not seen in the early stages. However, the early effects of PD are reflected in the speech of patients. The patients have soft or slurred speech with difficulty in utterances. More than 90% of patients have problems with speech [6]. Researchers in the past have leveraged this fact to build predictive models for the early detection of PD. With the improvement in computational ability and great research in the field of machine learning (ML), various diseases are diagnosed widely using different ML algorithms [7]. Various efforts are being made by the researchers to classify PD patients from healthy subjects using various machine learning techniques that leverages different early symptoms like speech deterioration, posture impairment, etc. to help build the intelligent systems for early diagnosis of AD.

As discussed earlier, there are certain symptoms that appear earlier than effects in the motor system. Kotsavasiloglou et al. [8] used simple drawing movement-based features for building a classification model for the delineation of PD patients from healthy subjects. The wrapper-based feature selection technique based on Naive Bayes (NB) algorithm gave an accuracy of 90.90%. Abdulhay et al. [9] used gait and tremor-based features for building an ML model. Features were selected using a peak detection algorithm and pulse duration to achieve an accuracy of 92.7%. Similarly, Yaman et al. [10] proposed a feature selection-based model using speech-based features. The model used Relief based feature selection method [11] that gave an accuracy of 91.25% with Support Vector Machines (SVM). The accuracy for cosine KNN was 0.912. Naranjo et al. [12] also proposed a two-stage variable selection approach for the classification of AD using speech-based features. The work considers intrasubject variability for variable selection and classification and was able to give an accuracy of 86.2% with sensitivity and specificity of 0.825 and 0.900 respectively. Over the past 15 years or so, great efforts are made in leveraging early symptoms of PD for easy and rapid diagnosis of the patients with PD. Various models based on speech, gait, posture, movement, etc. are developed. Since speech is one of the earliest symptoms of PD with noticeable changes in speech pattern, more emphasis has been laid on making models that extensively use speech-based features. In this paper, the machine learning method based on the feature selection technique has been used for the delineation of PD patients from healthy subjects on the basis of their speech-based features. Section 2 of the paper presents and describes the proposed methodology in detail. The experimental results from the proposed methodology are given in section 3 along with the comparisons to past works. Similarly, section 4 which is also the last section concludes the paper along with future works that can be done in the field.

II. METHODOLOGY

Fig. 1 shows the proposed model for the detection of the patients with Parkinson's disease from healthy subjects using feature selection technique and Twin Support Vector Machines (TSVM) [13]. The framework starts with collection of data needed for the research and goes through the preprocessing of collected data, selection of features, testing of different learning algorithms and evaluation of the performance of the models that have been trained and validated. The experiments were done with various learning algorithms to get an overview of the performance of TSVM as compared to traditional ML algorithms. For each of the learning algorithms, 10-fold cross-validation has been done for obtaining more consistent results. Each component of the given framework has been discussed in detail subsequent sections.

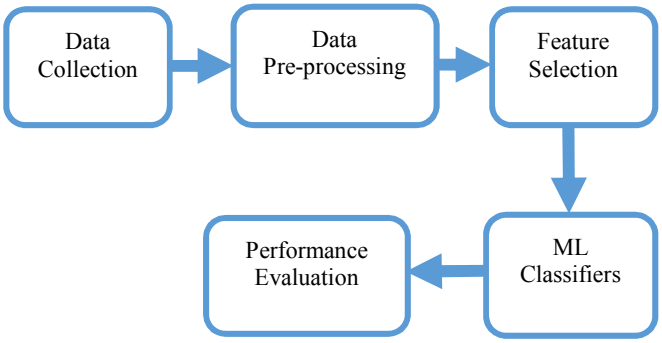


Fig. 1. Model for detection of Parkinson's Disease

A. Data Collection

The dataset that is used in this study was prepared by Little et al. [14] and is publicly available. The dataset consists of the speech-recordings of 31 male and female participants among which 23 of the participants were diagnosed with PD. The time since the first diagnosis of PD ranges from 0 to 28 years. There is an average of six voice-recordings per participant. The range of the length of voice measurement is 1 to 36 seconds. There are 147 recordings belonging to patients with PD and 48 recordings belonging to healthy subjects. The voice-recordings were extracted into 23 features as described in table 1.

TABLE I. FEATURES USED IN THE DATASET

S.N.	Attributes	Description
1.	MDVP:Fo(Hz)	Average Vocal Fundamental Frequency
2.	MDVP:Fhi(Hz)	Maximum Vocal Fundamental Frequency
3.	MDVP:Flo(Hz)	Minimum Vocal Fundamental Frequency
4.	MDVP:Jitter(%)	Several measures of variation in fundamental frequency
5.	MDVP:Jitter(Abs)	
6.	MDVP:RAP	
7.	MDVP:PPQ	
8.	Jitter:DDP	
9.	MDVP:Shimmer	Several measures of variation in amplitude
10.	MDVP:Shimmer(dB)	
11.	Shimmer:APQ3	
12.	Shimmer:APQ5	
13.	MDVP:APQ	
14.	Shimmer:DDA	
15.	NHR	Measures of the ratio of noise to tonal components in voice
16.	HNR	
17.	Status	0: Healthy Subjects and 1: PD Patients
18.	RPDE	Two non-linear dynamical complexity measures
19.	D2	
20.	DFA	Signal Fractal scaling exponent
21.	Spread 1	Three nonlinear measures of fundamental frequency variation
22.	Spread 2	
23.	PPE	

B. Data Pre-processing

The dataset provided by Little et al. [14] is an unbalanced data with 147 recordings belonging to PD class and 48 recordings that belong to healthy subjects. Imbalanced data gives poor performance in terms of sensitivity and specificity. For that reason, data has been balanced using Synthetic Minority Oversampling Technique (SMOTE) [15]. It balances the data by creating synthetic samples for the minority class and is one of the most widely used oversampling techniques. After applying SMOTE, the size of data becomes equal with 147 observations for each class. The dataset has features that are in very different ranges from one another. For that reason, data must be made on the same scale. Normalization reduces the impact of large valued features over small valued features. This helps learning algorithms to find the patterns that are affecting the data. In the experiment, the data were scaled and made in the range of -1 to 1 using MinMaxScaler of scikit-learn library [16]. The data were scaled according to formula in eq. (1).

$$x'' = 2 \times \frac{x - min(x)}{max(x) - min(x)} - 1 \quad (1)$$

C. Feature Selection

There are a lot of irrelevant or redundant features that are present in the data. Such redundant or irrelevant features do not increase the accuracy but might decrease the accuracy. The feature selection technique makes learning faster and makes the model more interpretable. In this experiment, correlation-based feature subset selection has been used. This technique selects the subset of features that are highly correlated with the class while having low intercorrelation [17]. Such features are searched using a forward selection method with Weka's BestFirst searching technique [18]. A total of 13 out of 23 features were selected with this feature selection technique. The features selected by the technique can be seen table II.

TABLE II. FEATURES SELECTED BY SELECTION TECHNIQUE

Rank	Attributes	Rank	Attributes
1.	MDVP:Fo(Hz)	8.	MDVP:APQ
2.	MDVP:Fhi(Hz)	9.	D2
3.	MDVP:Flo(Hz)	10.	DFA
4.	MDVP:Jitter(Abs)	11.	Spread 1
5.	MDVP:RAP	12.	Spread 2
6.	MDVP:Shimmer	13.	PPE
7.	Shimmer:APQ5		

D. Classifiers

In this experiment, six algorithms were used for training. Apart from TSVM [13], different machine learning classifiers were taken to draw a comparative analysis of the robustness of one model over the other models. In this experiment, logistic regression (LR), Support Vector Machines (SVM), Naive Bayes (NB), Decision Tree (DT) and K-Nearest Neighbor (KNN) were used. For all learning algorithms, 10-fold cross validation has been done and the average from performances in each fold of 10-fold cross validation is taken for the final evaluation. The classifiers used are briefly described below:

Logistic Regression: It is a basic binary classifier that uses logistic function to predict the class in a binary classification [19]. Unlike other classifiers like decision tree, which is based on different decision rules, it is based on simple probabilistic model.

Support Vector Machines: It is a discriminative classifier which is defined mathematically by separating hyperplane [24]. It is very efficient in binary classification with small data size. Since the data size of the data use in the experiment is comparatively smaller, SVM is expected to show good performance. The algorithm is used in solving regression problems also.

Naïve Bayes: Naïve Bayes, a classifier that belongs to the family of simple probabilistic classifier, is widely used in binary classifications [21]. Given the class variable, a Naïve Bayes classifier works with assumptions that all attributes are independent.

Decision Tree: It is one of the simplest machine learning algorithms that can be used to solve both classification and regression tasks. Decision tree, a simple flowchart-like tree structure, helps in making binary classification by making decisions based on chance event outcomes iteratively for different attributes [22].

K-Nearest Neighbor: K-Nearest Neighbor classifier, commonly known as KNN classifier makes the decision based on similarity measure [23]. It is a simple clustering algorithm that keeps on recording given observations along with its attributes to give prediction on test data based on those recorded instances.

Twin-Support Vector Machines: Twin Support Vector Machine (TSVM) is a machine learning algorithm that builds on the very popular support vector machine algorithm [13]. To reduce computational complexity and increase accuracy of the traditional SVM, many improved algorithms have been presented. One of them is generalized eigenvalue proximal support vector machines (GEPSVM), which constructs two non-parallel planes assuming that the data points of a certain class is very close to other members of the same class and is far from members of another class. The data points are clustered around a particular hyperplane and using this fact, the classification is done. GEPSVM has improved speed but its classification is low. Building on GEPSVM, Jayadeva et al. [13] proposed the twin support vector algorithm. Unlike SVM, TSVM generates two non-parallel planes in a manner where each plane is nearer to one of the given classes in binary classification whereas those two planes remain as far as possible from one another. This is done by solving two quadratic programming problems. Two non-parallel planes are obtained in which the data points corresponding to a particular class are clustered around a certain plane. Mathematical expressions involved in the Twin SVM algorithm as well as its brief explanation is given below.

Let $X = [X_1, X_2, X_3, \ldots, X_m]$ denote the matrix which contains the training example and let 'm' be the number of training examples and 'n' be the number of features in each training instance, $X_i = [X_{i1}, X_{i2}, \ldots, X_{in}]^T$, Also let $Y_i \in \{1, -1\}$ be the output classes to which the training example belongs to. Now we further partition X into $X^{(1)}$ and $X^{(2)}$, where $X^{(1)}$ contains those training examples which belong to $Y_i = 1$ and $X^{(2)}$ contains those training examples which belong to $Y_i = -1$. Let the number of training examples in $X^{(1)}$ and $X^{(2)}$ be m1 and m2 respectively. Therefore the size of the matrix $X^{(1)}$ and $X^{(2)}$ will be $(n \times m1)$ and $(n \times m2)$ respectively. $W^{(1)} = [W^{(1)}_1, W^{(1)}_2, \ldots, W^{(1)}_n]$, $b^{(1)}$, $W^{(2)} = [W^{(2)}_1, W^{(2)}_2, \ldots, W^{(2)}_n]$, $b^{(2)}$, q are the trainable parameters. We then solve the following two quadratic programming problems to obtain the classifying plane.

$$\underset{W^{(1)},b^{(1)},q}{Min} \ \frac{1}{2}[W^{(1)}X^{(1)} + b^{(1)}o_1][W^{(1)}X^{(1)} + b^{(1)}o_1]^T + c_1o_2q$$
$$\text{Subject to } -[W^{(1)}X^{(2)} + b^{(1)}o_2] + q \geq o_2, \ q \geq 0 \qquad (2)$$

And,

$$\underset{W^{(2)},b^{(2)},q}{Min} \ \frac{1}{2}[W^{(2)}X^{(2)} + b^{(2)}o_2][W^{(2)}X^{(2)} + b^{(2)}o_2]^T + c_2o_1q$$
$$\text{Subject to } [W^{(2)}X^{(1)} + b^{(2)}o_1] + q \geq o_1, \ q \geq 0 \qquad (3)$$

Where $c_1, c_2 \geq 0$, are hyper parameters and o_1, o_2 are matrices of ones of appropriate dimensions.

The algorithm finds two hyperplanes

$$W^{(1)}X + b^{(1)} = 0 \qquad (4)$$
$$W^{(2)}X + b^{(2)} = 0 \qquad (5)$$

If a new instance X is closer to plane (3) then it is classified as belonging to Y = 1 and if the new instance X is closer to plane (4), then it is classified as belonging to Y = -1. The twin SVM algorithm is approximately four times as fast as the SVM algorithm. Let $T_{SVM}(m)$ be the time complexity of the SVM algorithm and let $T_{TSVM}(m)$ be the time complexity of the Twin SVM algorithm, where m is the number of training examples. Then we know that $T_{SVM}(m) = m^3$. Assuming that $m1 = m2 = \frac{m}{2}$, $T_{TSVM}(m) = 2 \times (\frac{m}{2})^3 = \frac{m^3}{4} = \frac{T_{SVM}(m)}{4}$. Hence these equations show that the time taken to solve Twin SVM is 4 times as fast as the time taken to solve SVM.

The above equation finds a linear hyperplane but the Twin SVM, similar to SVM can be modified using kernels to give nonlinear surfaces as well so that the algorithm finds nonlinear decision boundaries. We modify the quadratic programming equations to solve for the following planes.

$$V^{(1)}K(X, C) + b^{(1)} = 0 \qquad (6)$$
$$V^{(2)}K(X, C) + b^{(2)} = 0 \qquad (7)$$

Where $V^{(1)}$, $V^{(2)}$, $b^{(1)}$, $b^{(2)}$ are trainable parameters and K defines the kernel function that is used and $C = [X^{(1)} X^{(2)}]$

The quadratic programming problems are then formulated as:

$$\underset{V^{(1)},b^{(1)},q}{Min} \ \frac{1}{2} \| V^{(1)}K(X^{(1)}, C) + b^{(1)}o_1 \|^2 + c_1o_2q$$
$$\text{Subject to } -[V^{(1)}K(X^{(2)}, C) + b^{(1)}o_2] + q \geq o_2, q \geq 0 \qquad (8)$$

And,

$$\underset{V^{(2)},b^{(2)},q}{Min} \ \frac{1}{2} \| V^{(2)}K(X^{(2)}, C) + b^{(2)}o_2 \|^2 + c_2o_1q$$
$$\text{Subject to } [V^{(2)}K(X^{(1)}, C) + b^{(2)}o_1] + q \geq o_1, q \geq 0 \qquad (9)$$

E. Performance Evaluation

The performance of all the classifiers used above is evaluated on the basis of accuracy, precision, recall, f-score, area under the curve (AUC), sensitivity, and specificity respectively as shown in Eqs. (10)-(15). The formulae for given equations are based on confusion matrix as shown in table III.

TABLE III. CONFUSION MATRIX FOR THE EXPERIMENT

Prediction → Actual Data ↓	Classified as PD patients	Classified as Healthy Subjects
Real Patients with PD	True Positives (TP)	False Negatives (FN)
Real Healthy Subjects	False Positives (FP)	True Negatives (TN)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (14)$$

$$Specificity = \frac{TN}{TN + FP} \quad (15)$$

Where, TP, TN, FP, and FN are True Positives, True Negatives, False Positives and False Negatives, respectively.

III. RESULTS AND DISCUSSION

The selected features using the feature selection technique discussed above are shown in table II. Table IV shows the performance of different classifiers with and without feature selection technique. Performance measures like Precision, Recall, F-score are shown in table IV. Similarly, table V shows the performance measures like accuracy (acc), AUC, sensitivity, and specificity.

Fig. 2 shows the graphical representation for accuracy scores with and without feature selection techniques. From the results in table IV and V, it can be observed that with selection of features, there is increase in the classification accuracy as well as other performance measures. This shows

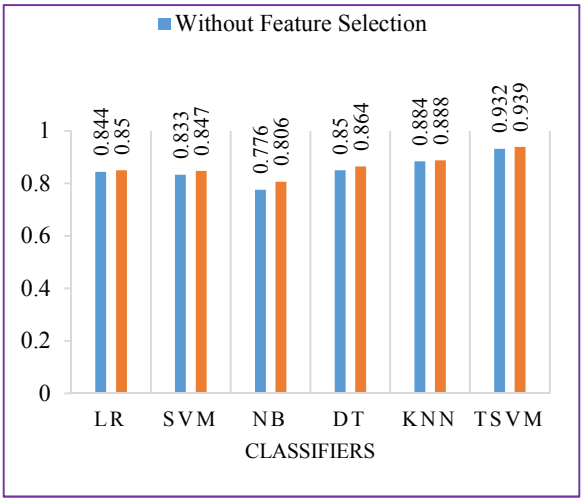


Fig. 2. Classification Accuracy of various Classifiers

that there are some redundant and irrelevant features that are removed by selecting just the relevant features needed for the classification problem. The feature selection provides optimal decisional space for the classification algorithms. When the performance of all six classifiers are compared, TSVM is the best classifier with an accuracy of 0.932 without feature selection and 0.939 with feature selection. In the case of TSVM also, as general, accuracy has been increased due to feature selection algorithm.

From table IV and V, it can be further observed that the twin support vector machines with feature selection is more robust in terms of performance measures like precision, recall, f-score, and AUC. The precision, recall, and f-score for the TSVM with feature selection have same value of 0.940 whereas without feature selection, precision, recall and f-score is 0.930 each. Similarly, the values of AUC with and without feature selection for TSVM are 0.932 and 0.939 respectively. There is a slight increase in AUC value for TSVM using feature selection. It is interesting to note that there is a great jump in AUC for other classifiers also when feature selection technique is used.

Table VI shows a comparison of the proposed methodology with past works done in the field. It is interesting to note that TSVM is the best performing algorithm that performs better than other algorithms used for the classification of PD from healthy subjects. Among the various modes of features like gait and tremor based, handwriting based and speech-based modes, the presented TSVM algorithm with speech-based features performs the best. It shows that the TSVM algorithm with the feature selection technique used in the experiment undoubtedly performs very well. Further optimization and hyperparameter tuning can yield even better results [20].

TABLE IV. PRECISION, RECALL AND F-SCORE WITH DIFFERENT CLASSIFIERS

Classifier	Without Feature Selection			With Feature Selection		
	Precision	Recall	F-score	Precision	Recall	F-score
Logistic Regression	0.844	0.844	0.844	0.850	0.850	0.850
SVM	0.842	0.833	0.832	0.856	0.847	0.846
Naïve Bayes	0.803	0.776	0.770	0.820	0.806	0.804
Decision Tree	0.860	0.850	0.850	0.880	0.880	0.880
KNN	0.900	0.880	0.880	0.900	0.890	0.890
TSVM	0.930	0.930	0.930	0.940	0.940	0.940

TABLE V. ACCURACY, AUC, SENSITIVITY, SPECIFICITY WITH DIFFERENT CLASSIFIERS

	Without Feature Selection				With Feature Selection			
Classifier	Accuracy	AUC	Sensitivity	Specificity	Accuracy	AUC	Sensitivity	Specificity
Logistic Regression	0.844	0.929	0.837	0.850	0.850	0.935	0.844	0.857
SVM	0.833	0.833	0.755	0.912	0.847	0.847	0.769	0.925
Naïve Bayes	0.776	0.871	0.626	0.925	0.806	0.804	0.701	0.912
Decision Tree	0.850	0.850	0.790	0.910	0.864	0.880	0.840	0.920
KNN	0.884	0.884	0.800	0.970	0.888	0.890	0.800	0.980
TSVM	0.932	0.932	0.910	0.970	0.939	0.939	0.880	0.990

TABLE VI. COMPARISION OF THE PROPOSED METHODOLOGY WITH PAST WORKS

Papers	Algorithm	No. of features	Feature	Feature Selection	Acc.	Precision	Recall	F-measure
Abdulhay et al. [9]	SVM	4	Gait and tremor based	Peak detection and pulse duration	0.927	0.944	0.974	0.954
Kotsavasiloglou et al. [8]	NB	9	Hand-writing Based	Wrapper Based	0.909	-	-	-
Yaman et al. [10]	Cosine KNN	66	Speech-based	Relief based feature selection	0.9123	0.9132	0.9125	0.912
	Gaussian SVM				0.9125	0.9125	0.9125	0.9125
Naranjo et al. [12]	Two stage variable selection and classification	44	Speech-Based	Two stage variable selection Bayesian Approach	0.862	0.891	0.825	0.857
Proposed Methodology	TSVM	13	Speech-Based	Correlation Based	0.939	0.940	0.940	0.940

IV. CONCLUSION AND FUTURE WORKS

From the experiments done, it can be deciphered that TSVM can prove to be a better classifier for binary classification problem like delineation of Parkinson's disease and healthy subjects. The feature selection algorithm further improves the accuracy by significant figure. More techniques on features selection can be explored. The experiments were done only on speech modality, there can be use of multimodal methods for building classification models. Speech modality can be added with handwriting-based features or gait-based features to get better performance from the model. The model also becomes more interpretable when multimodal techniques are used. The search for biomarkers and early symptoms is undoubtedly of high importance. Apart from that, better machine learning algorithms and deep learning algorithms should be used to provide better decisional space for making more intelligent predictions. Similarly, various optimization techniques can enhance the classification performance of the classifiers which is the scope for future works.

REFERENCES

[1] T. Sampson, "The impact of indigenous microbes on Parkinson's disease," Neurobiology of disease, vol. 135, p. 104426, 2020.

[2] W. Dauer and S. Przedborski, "Parkinson's disease: mechanisms and models," Neuron, vol. 39, no. 6, pp. 889-909, 2003.

[3] (10 May). Parkinson's Disease. Available: www.nhs.uk/conditions/parkinsons-disease

[4] A. E. Lang and A. M. Lozano, "Parkinson's disease," New England Journal of Medicine, vol. 339, no. 16, pp. 1130-1143, 1998.

[5] J. L. Cummings, "Depression and Parkinson's disease: a review," The American journal of psychiatry, 1992.

[6] C. O. Sakar and O. Kursun, "Telediagnosis of Parkinson's disease using measurements of dysphonia," Journal of medical systems, vol. 34, no. 4, pp. 591-599, 2010.

[7] A. Ghimire, S. Thapa, A. K. Jha, A. Kumar, A. Kumar, and S. Adhikari, "AI and IoT Solutions for Tackling COVID-19 Pandemic", in 2020 International Conference on Electronics, Communication and Aerospace Technology, 2020: IEEE

[8] C. Kotsavasiloglou, N. Kostikis, D. Hristu-Varsakelis, and M. Arnaoutoglou, "Machine learning-based classification of simple drawing movements in Parkinson's disease," Biomedical Signal Processing and Control, vol. 31, pp. 174-180, 2017.

[9] E. Abdulhay, N. Arunkumar, K. Narasimhan, E. Vellaiappan, and V. Venkatraman, "Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease," Future Generation Computer Systems, vol. 83, pp. 366-373, 2018.

[10] O. Yaman, F. Ertam, and T. Tuncer, "Automated Parkinson's disease recognition based on statistical pooling method using acoustic features," Medical Hypotheses, vol. 135, p. 109483, 2020.

[11] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," Journal of biomedical informatics, vol. 85, pp. 189-203, 2018.

[12] L. Naranjo, C. J. Perez, J. Martin, and Y. Campos-Roca, "A two-stage variable selection and classification approach for Parkinson's disease detection by using voice recording replications," Computer methods and programs in biomedicine, vol. 142, pp. 147-156, 2017.

[13] Jayadeva, R. Khemchandani and S. Chandra, "Twin support vector machines for pattern classification," IEEE Transactions on pattern analysis and machine intelligence, vol. 29, no. 5, pp. 905-910, 2007.

[14] M. Little, P. McSharry, E. Hunter, J. Spielman, and L. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," Nature Precedings, pp. 1-1, 2008.

[15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," Journal of artificial intelligence research, vol. 16, pp. 321-357, 2002.

[16] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," Journal of machine Learning research, vol. 12, pp. 2825-2830, 2011.

[17] M. A. Hall, "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning," in Proceedings of the Seventeenth International Conference on Machine Learning, 2000, pp. 359-366: Morgan Kaufmann Publishers Inc.

[18] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," ACM SIGKDD explorations newsletter, vol. 11, no. 1, pp. 10-18, 2009.

[19] D. G. Kleinbaum, K. Dietz, M. Gail, M. Klein, and M. Klein, Logistic regression. Springer, 2002.

[20] N. Bansal, R. Gautam, R. Tiwari, S. Thapa, and A. Singh, "Economic Load Dispatch Using Intelligent Particle Swarm Optimization", in 2020 2nd International Conference on Intelligent Computing, Information and Control Systems (ICICCS 2020), 2020: Springer.

[21] I. Rish, "An empirical study of the naive Bayes classifier," in IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001, vol. 3, no. 22, pp. 41-46.

[22] S. Thapa, P. Singh, D. K. Jain, N. Bharill, A. Gupta, and M. Prasad, "Data-Driven Approach based on Feature Selection Technique for Early Diagnosis of Alzheimer's Disease", in 2020 International Joint Conference on Neural Networks (IJCNN), 2020: IEEE, pp. 1-8.

[23] A. Saxena et al., "A review of clustering techniques and developments," vol. 267, pp. 664-681, 2017.

[24] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273-297, 1995.

View publication stats

Indian Journal of Pure & Applied Physics
Vol. 58, November 2020, pp. 804-811

Force and temperature analysis during distinct machining environment using an optimization approach

Anurag Sharma[a], R C Singh[b], Ranganath M Singari[b*] & S L Bhandarkar[c]

[a]Department of Mechanical Engineering, Delhi Technological University, Delhi-110 042 India

[b]Department of Mechanical Engineering, Delhi Technological University, Delhi-110 042 India

[c]Board of Technical Education, Muni Maya Ram Marg, Pitampura, Delhi- 110 088 India

Received 18 January 2020; accepted 8 October 2020

Cryogenic cooling with liquid nitrogen has a wide field of scientific and engineering applications. The aim of the presented research work is to analyze the impact on cutting force and temperature during the LN_2 single supply at (i) rake face and double supply (ii) rake and flank face of TiN coated carbide cutting tool during turning of AISI D3 steel. Taguchi based S/N ratio was used for Design of Experiments (DoE) of L_{18} orthogonal array. The four control factors were selected. The three factors like speed, feed and depth of cut were varied to three levels. The fourth factor was the machining condition which was varied to two levels. The cutting force and temperature were declined by 32-48% and 34-46% during the double supply of LN_2 at rake and flank face as compared to single supply at rake face. The influence and percentage of contribution of each factor were found out by Analysis of Variance (ANOVA). The double supply method with LN_2 supply at rake and flank face had the highest effect of contribution in the percentage as 77.13% and 87.42% for cutting force and temperature respectively. The values obtained during confirmatory runs were very close to the optimized results within the limits.

Keywords Liquid nitrogen, cutting force, temperature, optimization, ANOVA

1 Introduction

The conventional cutting fluids may not be able to provide sufficient cooling for difficult to cut materials. The process may create problems like splashing, health-related problems for an operator like nausea, itching to eyes and skin infections etc. The dry machining condition could be environmental friendly but the manufacturing rate could be low due to operating at slow cutting parameters[1-5]. The green and clean machining may be supported by using cryogenic cooling with LN_2. The health-related problems may be solved with cleanliness and neatness in the machining process. Liquid nitrogen is non-toxic, odourless and tasteless. The heat removing rate is highly efficient. The boiling point of liquid nitrogen is -196°C[6-9]. Gupta *et al.*[10] performed experiments during dry and cryogenic cooling with LN_2 on AISI 1040 steel. It was found that tool wear and cutting force was reduced to 55.45-65.55% and 61.94-96.6% during cryogenic cooling with LN_2 as compared with dry turning. Shokarani *et al.*[11] used cobalt chromium for machining with LN_2, MQL (minimum quantity lubrication) and flood. It has been found on comparing with LN_2 surface roughness was reduced by 35% and 42% as compared with MQL and flood machining. Flank wear during LN_2 was lowered by 26 times and 17 times as compared with MQL and flood cooling condition. Sartori *et al.*[12] investigated surface morphology of titanium alloy Ti 6Al 4V machine during different machining conditions like dry, wet, MQL, LN_2, CO_2, LN_2(rake)+MQL, LN_2(flank)+ MQL(rake) and CO_2(rake)+MQL(flank). The minimum surface roughness was found during cryogenic cooling with LN_2. Dhanachezian *et al.*[13] used duplex stainless steel 2205 as work piece for analysing the influence of LN_2 during turning. It was found that cutting temperature and cutting force were declined by 53-58% and 30-43% during cryogenic cooling with LN_2 on comparing with dry machining. Sharma *et al.* [14] used cold working die steel AISI D3 as workpiece with TiN coated carbide tool for analysing the effect of LN_2 on machiniability characteristics. It was found that surface roughness, cutting force, flank wear length and temperature were declined by 15-20%, 51-55%, 35-59% and 23-52% respectively on comparing with dry machining condition. The machining operation consists of different machining parameters. The optimization process gives a better combination of parameters for

*Corresponding author: (E-mail id: ranganath@dce.ac.in)

achieving the low (surface roughness, tool wear, cutting force, temperature *etc.*) or high machining response (material removal rate). The resources are used without any wastage. Taguchi Technique was discovered by Genichi Taguchi in 1950. Design of experiments based on Taguchi S/N ratio eliminates unnecessary hit and trial of experiments which reduce the number of experiments to be performed to a definite value. This saves costs associated with workpiece material, tool material, power supply, energy consumption etc. The process is time and economy saving[14, 15]. The optimization process has scientific and engineering applications which are helpful in the design process [16,17]. Gupta *et al.*[18] used Taguchi based optimization in determining the selection of process among dry, wet and LN_2. It was found that cryogenic cooling with LN_2 gave optimum value for surface roughness and cutting force. The different types of steel like EN 47, EN 19, AISI 1045, titanium alloys *etc.* were used by researchers and optimized for machinability characteristics [19-22]. High chromium tool steel alloy AISI D3 (HRC 60) is used in manufacturing blanking & forming dies, forming tools, press tools, punches, bushes and wear-resistant moulds. It is categorised as difficult to machine materials. Cutting parameters for turning of AISI D3 steel alloy is speed 30-260 m/minute, feed 0.05-0.15 mm/revolution and depth of cut 0.2-0.5 mm[14]. The exhaustive literature review shows the combination of AISI D3 workpiece and TiN coated carbide cutting tool is not used during cryogenic turning. in single supply on (i) rake face and double supply at (ii) rake and flank face. In the presented research work AISI D3 was used as workpiece material and TiN coated carbide coated cutting inserts as the cutting tool. The experiments were performed on lathe machine. DoE was Taguchi based on L_{18} [OA]. Two levels of machining condition were proposed. In one condition LN_2 was directly supplied at the rake face of cutting insert and in another condition LN_2 was directly supplied at rake and flank face. The response values of cutting force and temperature were analysed and optimized. The confirmatory runs were performed for checking the response values was in the predicted limit or not.

2 Experimental Processes

2.1 Materials

Round bars of diameter 45 mm and 300 mm long of AISI D3 steel was used in the performance of the experiment. The chemical analysis of workpiece material in the percentage of weight as C(2.01), Si(0.249), Mn (0.426), S(0.018), P(0.021), Cr(11.03), Ni(0.075), Mo(0.08), Co(0.012), Nb(0.22), V(0.037), W(0.088) and Fe (85.929). Carbide cutting inserts with coated TiN in diamond shape were used during performing experiments. The geometry of cutting insert is in accordance with specification DCMT11T308HQ grade PV 20. The thickness is 4mm, end relief angle 11º, front rake angle 0º, nose radius 8 mm. The elements were found as percentage of weight in cutting insert were C(7.06), Co(16.27), Mn(0.29), V(2.06), Nb(5.24), Ni(4.94), W(15.84), Mo(0.67) and Ti(47.64).

2.2 Machining setup

The turning experiments were performed on a lathe machine. The distinct machining environment was created by the direct supply of LN_2 from Dewar container A 55. Uniform flow of air was maintained by an air compressor with a pressure regulator connected to the container. Insulated pipes were connected from container to rake face of cutting insert and in another approach to flank and rake face. Piezo-electric type dynamometer was used for cutting force measurement. Digital infrared non-contact type thermometer was used to measure temperature by the end of the experiment. Every stage of the experiment was started with unused new cutting insert.

2.3 Design of Experiment (DoE)

Turning parameters were selected during the performance of experiments are shown in Table 1. One parameter was related to two distinct machining conditions. Other three parameters were related to 3V', F' and D'. Each was selected with three values in increasing order. Taguchi based orthogonal array [OA] L_{18} saved energy, raw material and operating cost. This may be due to declining the number of experiments for more than 50%.

3 Results

3.1 Probability plot investigations

Probability plots of each response are shown in Fig. 1. The data values were approximately aligned to

Table 1 — Turning parameters during the performance of experiments

Levels	Parameters			
	Mc'	V' (m/min.)	F'(mm/rev.)	D'(mm)
Level 1	LN_2-SJ	50	0.06	0.30
Level 2	LN_2-DJ	70	0.08	0.35
Level 3	-	90	0.10	0.40

the straight central line and normally distributed. The responses were validated by probability plots. This confirms for using data for further analysis.

3.2 Optimization based on Taguchi (Signal to Noise ratio)

Taguchi (S/N) based optimization was used to control the number of experiments to arrive at the targeted outcome. Eq. 1 shows the smaller the better characteristic of continuous response function depending on the optimization principle.

Smaller is the better characteristic,

$$\frac{s}{n} = -10log\frac{1}{n}(\sum x^2) \quad \ldots (1)$$

The values of turning performance characteristics & S/N ratio of turning performance characteristics are shown in Table 2. ANOVA means analysis of variance was used for completing and signifying the importance of process parameters contributing factor. The p-value signifies the level of 5% (confidence level of 95%) for all responses.

3.3 Cutting force

It is shown in Fig. 2 that cutting force was more in experiments 1-9 with a directly supplied of LN_2 at (rake face only) than experiments 10-19 with a

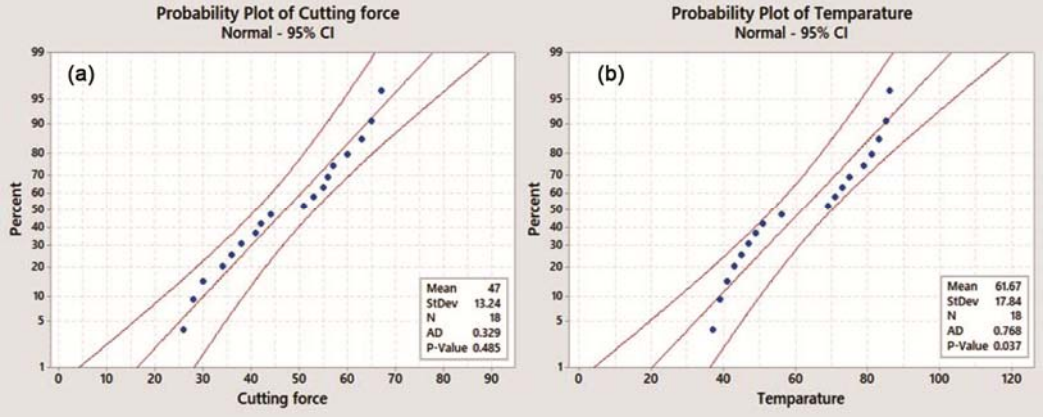


Fig. 1 — Probability plots for (a) cutting forces and (b) temperature

Table 2 — L_{18} Taguchi [OA] DoE with turning parameters and values of each response

Run	Turning parameters				Values of each response		S/N value of each response	
	Machining condition	V' (m/min.)	F' (mm/rev.)	D' (mm)	CF'	TR'	CF'	TR'
1	LN_2-SJ	50	0.06	0.30	67	69	-36.5215	-36.777
2	LN_2-SJ	50	0.08	0.35	65	71	-36.2583	-37.0252
3	LN_2-SJ	50	0.10	0.40	63	73	-35.9868	-37.2665
4	LN_2-SJ	70	0.06	0.30	56	75	-34.9638	-37.5012
5	LN_2-SJ	70	0.08	0.35	57	79	-35.1175	-37.9525
6	LN_2-SJ	70	0.10	0.40	60	81	-35.563	-38.1697
7	LN_2-SJ	90	0.06	0.35	51	83	-34.1514	-38.3816
8	LN_2-SJ	90	0.08	0.40	53	85	-34.4855	-38.5884
9	LN_2-SJ	90	0.10	0.30	55	86	-34.8073	-38.69
10	LN_2-DJ	50	0.06	0.40	44	37	-32.8691	-31.364
11	LN_2-DJ	50	0.08	0.30	41	39	-32.2557	-31.8213
12	LN_2-DJ	50	0.10	0.35	42	41	-32.465	-32.2557
13	LN_2-DJ	70	0.06	0.35	34	43	-30.6296	-32.6694
14	LN_2-DJ	70	0.08	0.40	36	45	-31.1261	-33.0643
15	LN_2-DJ	70	0.10	0.30	38	47	-31.5957	-33.442
16	LN_2-DJ	90	0.06	0.40	26	56	-28.2995	-34.9638
17	LN_2-DJ	90	0.08	0.30	28	51	-28.9432	-34.1514
18	LN_2-DJ	90	0.10	0.35	30	49	-29.5424	-33.8039

directly supplied LN_2 at (rake and flank face) simultaneously. This may be due to better lubrication effect generated by LN_2 at flank and rake face. This reduced the magnitude of cutting force between the cutting tool and workpiece.

It is shown in Fig. 3(b) that variation of average CF' that on increasing cutting speed V', cutting force declined and on incrementing feed F' and depth of cut D' cutting force increased. Table 4 shows the respective values of CF' with turning parameters. Variation of average of S/N CF' is shown in Fig. 3(a). Cutting force was optimized at Mc' (level 2 = LN_2-DS), V' (level3 = 90m/min.), F' (level1 = 0.06) and D' (level2 = 0.35mm) as shown in Table 3. The predicted value was calculated from Eqs. 2 & 3. The predicted value of cutting force at an optimum level was 30.93N. It is shown in ANOVA Table 5 that Mc'(machining condition) was highly effective with the percentage of contribution (77.13%). Next followed by V', F' and D' in consecutive decreasing order as 18.74%, 0.48% and 0.07% respectively. F-value depicted relative importance in increasing order of Mc'(machining condition), V'(cutting speed), F'(feed) and D'(depth of cut). p-value was significant for Mc'(machining condition) and V' (cutting speed).

$$\eta_p = \overline{\eta}_p + \left(\overline{Mc_o} - \overline{\eta}_p\right) + \left(\overline{\alpha_o} - \overline{\eta}_p\right) + \left(\overline{\beta_o} - \overline{\eta}_p\right) + \left(\overline{\gamma_p} - \overline{\eta}_p\right) \quad \ldots (2)$$

$$Z_p = 10^{-\frac{\eta_p}{20}} \quad \ldots (3)$$

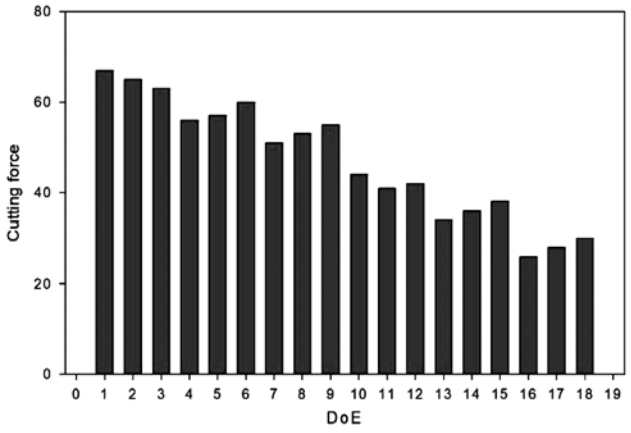


Fig. 2 — Experimental CF' according to DoE

3.4 Temperature

It is shown in Fig. 4 that temperature measured at the interface of cutting tool and workpiece was more in experiments 1-9 with directly supplied of LN_2 at (rake face only) than experiments 10-19 with a directly supplied of LN_2 at (rake and flank face). This may be due to the high coefficient of heat transfer. The double supply of LN_2 has incremented the faster rate of cooling of the cutting tool. The temperature lowered down and provided betters cooling at the interface of the cutting tool.

It is shown in Fig. 5(b) variation of mean TR' that on increasing cutting speed V' and feed F' was increased and on incrementing depth of cut D' temperature was marginally decreased than increased and respective values are shown in Table 4. Variation of mean S/N, TR' is shown in Fig, 5(a). Temperature was optimized at Mc' (LN2-DS), V' (50m/min.), F' (0.06 mm/rev.) and D' (0.35 mm) as shown in Table 3. The predicted value of temperature at an optimum level by using Eqs. 2 & 3 was 38.99°C. It was shown in ANOVA Table 5 that Mc'(machining condition)

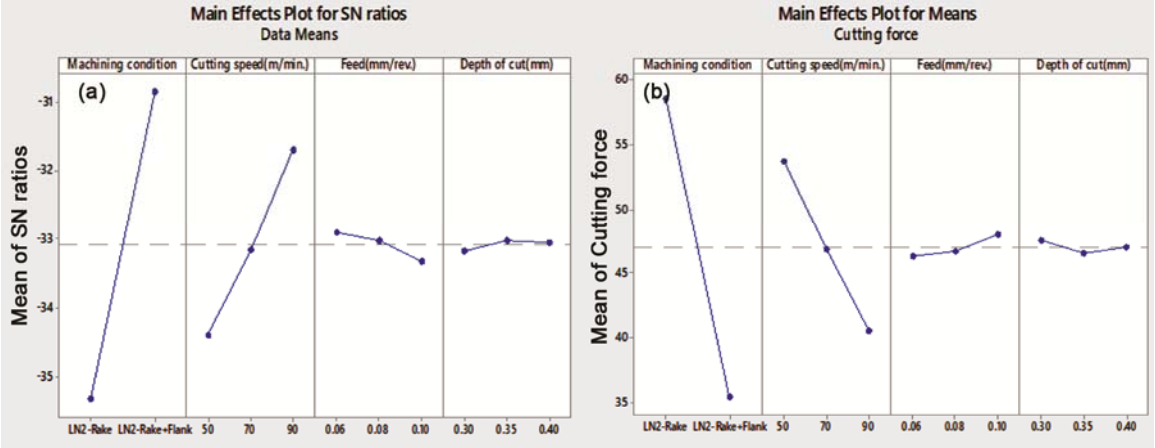


Fig. 3 — (a) Variation of average S/N CF' (b) variation of average CF' with different factors

Table 3 — Turning performance characteristics table for S/N ratio Cutting force (CF') and temperature (TR')

Turning performance characteristics	Level	Machining Condition	V'(m/min.)	F'(mm/rev.)	D'(mm)
CF'	1	-35.32	-34.39	**-32.91**	-33.18
	2	**-30.86**	-33.17	-33.03	**-33.03**
	3	-	**-31.70**	-33.33	-33.05
	Delta	4.46	2.69	0.42	-33.05
	Rank	1	2	3	4
TR'	1	-37.82	**-34.42**	**-35.28**	-35.40
	2	**-33.06**	-35.47	-35.43	**-35.35**
	3	-	-36.43	-35.60	-35.57
	Delta	4.76	2.01	0.33	0.22
	Rank	1	2	3	4

Table 4 — Turning performance characteristics table for means Cutting force (CF') and temperature (TR')

Turning performance characteristics	Level	Machining Condition Mc'	V'(m/min.)	F'(mm/ rev.)	D'(mm)
CF'	1	58.56	53.67	**46.33**	47.50
	2	**35.44**	46.83	46.67	**46.50**
	3	-	**40.50**	48.00	47.00
	Delta	23.11	13.17	1.67	1.00
	Rank	1	2	3	4
TR'	1	78.00	**55.00**	**60.50**	61.17
	2	**45.33**	61.67	61.67	**61.00**
	3	-	68.33	62.83	62.83
	Delta	32.67	13.33	2.33	1.83
	Rank	1	2	3	4

was highly effective with the percentage of contribution (87.42%). Next followed by V', F' and D' in consecutive decreasing order as 10.43%, 0.278% and 0.13% respectively. p-value was significant for Mc'(machining condition) and V'(cutting speed).

3.5 Confirmation experiment

Confirmation experiments were performed to check the closeness between predicted values and results during the performance of the experiment at an optimized level. (CI) confidence level was calculated from Eqs. 4 & 5[23]. The range has predicted that values obtained during the performance of the experiment are acceptable or not. The values were in good agreement with the predicted range were acceptable at the optimized level. The reliability of the condition was assumed to be 95%.

$$CI = \sqrt{\left[F(\alpha, 1, fe) Ve \left\{\left(\frac{1}{Ne}\right) + \left(\frac{1}{R}\right)\right\}\right]} \quad \ldots (4)$$

$$Ne = \frac{N}{(1+Td)} \quad \ldots (5)$$

Where, $F(\alpha,1,f_e)$ is the F-ratio required for 100(1- α) percent confidence level, fe is DOF for error =10, Ve= AdjMs for error, N = total number of experiment (18), R = number of replications for confirmation of experiments (0) and Td = total degree of freedom associated with mean optimum is (7). From standard statistical table, the value of F ratio for α = 0.05 is F(0.05,1,10) = 4.96. Substituting values from ANOVA Table 6 with the respective responses. CI value of cutting force (CF') was ±0.96N and temperature (T') was ±0.67

3.6 Confirmation tests

Confirmation tests were performed for specified cutting parameters and machining condition to evaluate the degree of accuracy between predicted and experimental values at optimal levels. It has been found that the experimental values of each response were within close range of limits with predicted values. Actual value of cutting force and temperature after performing experiment under optimized conditions were 31.0N and 38.99°C . This was within predicted range of ±0.96N and ±0.67°C.

3.7 Discussion

The experiments were performed on lathe machine using LN_2. The design of experiments was based on Taguchi S/N signal to noise ratio. This has reduced the number of experiments to more than 50% of orthogonal array of L_{18}. The liquid nitrogen reduced the cutting force and temperature during double supply at rake and flank face by 55-66% and 53-57% respectively on comparing with a single supply of LN_2 at rake face. The double supply of liquid nitrogen

incremented the efficiency in lowering down the temperature. The highly pressurised flow of LN_2 removed the fine debris generated at the interface of cutting insert and workpiece away from the machined surface which reduced the cutting force. The effect of thermal softening was eliminated. This provided the structural stability of the cutting insert. The tool became strong and retained the hardness. The cutting ability increased. The tool life increased. The optimized values were checked for confirmatory runs. It has been found that optimized response values were within the specified range of values. This proved that the performance of experiments was at the 95% confidence of level. The human-related problems like skin infection, nausea etc. were not found during the entire performance of experiments. The process was

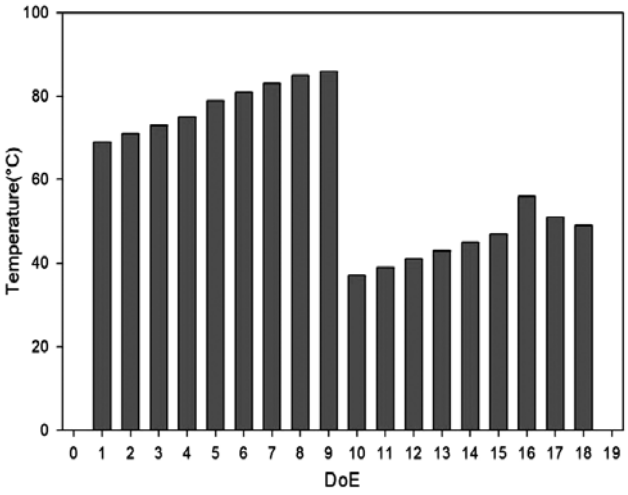


Fig. 4 — Experimental TR ' according to DoE

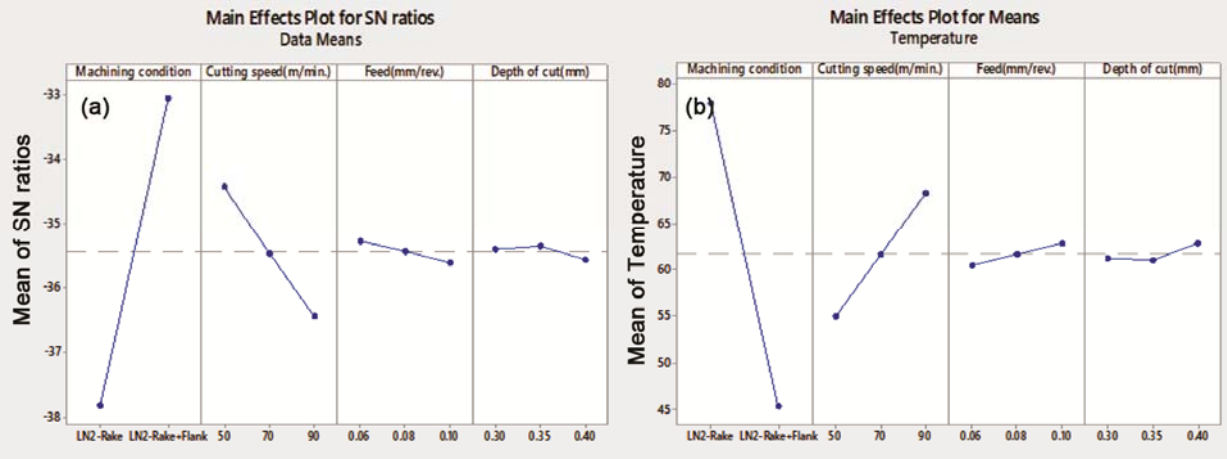


Fig. 5 — (a) Variation of average S/N TR' (b) variation of average TR' with different factors

Table 5 — Analysis of variance for means of Cutting force (CF') and Temperature (TR')

Turning performance characteristics	Source	DF	Adj SS	Adj MS	F-Value	P-Value	% Cont.
CF'	Machining Condition	1	89.463	89.4630	215.96	0.000	77.13
	V'(m/min.)	2	21.728	10.8642	26.23	0.000	18.74
	F'(mm/rev.)	2	0.561	0.2803	0.68	0.530	0.48
	D' (mm)	2	0.081	0.0403	0.10	0.908	0.07%
	Error	10	4.143	0.4143	-	-	-
	Total	17	115.975	-	-	-	-
TR'	Machining Condition	1	101.846	101.846	503.53	0.000	87.42
	V'(m/min.)	2	12.146	6.073	30.03	0.000	10.43
	F'(mm/rev.)	2	0.324	0.162	0.80	0.476	0.278
	D'(mm)	2	0.162	0.081	0.40	0.680	0.13
	Error	10	2.023	0.202	-	-	-
	Total	17	116.502	-	-	-	-

neat and green. The chips generated could be readily used as direct scrape for raw material. This further reduced the cost related to the handling of debris and chips. The process can be considered for sustainable green production.

Conclusions

Taguchi based L_{18} DoE was used in the presented analysing research work. The two machining conditions were selected. One is with LN_2 at (i) rake face and second was with LN_2 at (ii) rake and flank face. The responses like cutting force and temperature were optimized and check for accuracy by performing confirmation tests. The following conclusions have been made on the basis of above-mentioned research work.

1. LN_2 supply at rake and flank faces was significant and gave a better cooling than LN_2 supply at rake face.
2. LN_2 absorbed heat and evaporated quickly which formed a fluid layer between the cutting tool and workpiece. Dual supply (rake and flank face) reduced cutting force as compared to a single supply (rake face).
3. LN_2 reduced the temperature due to the effective and efficient property of heat-carrying capacity. Dual supply (rake and flank face) reduced temperature as compared to a single supply (rake face).
4. Cutting force was optimized at MC' (-DJ), V' (50 m/min.), F' (0.06 mm/rev.) D' (0.30 mm). The predicted value was 30.93N.
5. Machining condition with LN_2 supply at flank and rake face simultaneously has the highest contribution the percentage (77.13%). The value of cutting force after confirmation test is 31 N. The CI range limit was $\pm$ 0.96.
6. The temperature at the interface of cutting insert was optimized at MC' (LN_2-DJ), V' (50 m/min.), F' (0.06 mm/rev.) D' (0.35 mm). The predicted value was 38.99 °C.
7. Machining condition with LN_2 supply at flank and rake face simultaneously has the highest contribution the percentage (87.42%). The value of cutting force after confirmation test was 40 °C. The CI range limit was $\pm$0.67.
8. LN_2 was efficient and eco-friendly. This may be suitable for green machining

List of Symbols

V' : Cutting speed (m/min.)
F' : Feed (mm/rev.)
D' : Depth of cut (mm)
CF' : Cutting force (N)
TR' : Temperature (°C)
Mc' : Machining condition
LN_2-SJ : LN_2 at Rake Face
LN_2-DJ: LN_2 at Rake and Flank Face
LN_2 : Liquid Nitrogen,
DoE : Design of experiments
η_p : S/N ratio calculated at optimized parameters
$\overline{\eta}_p$: mean S/N ratio for all variables at optimized parameters
$\overline{Mc_o}$: mean S/N ratio when machining condition at optimized parameters
$\overline{\alpha}_p$: mean S/N ratio when cutting speed at optimized parameters
$\overline{\beta}_p$: mean S/N ratio when feed at optimized parameters
$\overline{\gamma}_p$: mean S/N ratio when depth of cut at optimized parameters

Acknowledgements

Authors are highly thankful to workshop and laboratories facilities shared by Delhi Technological University and Indian Institute of Technology, Delhi (India)

References

1 Byrne G & Scholta E, *CIRP Ann Manuf Technol*, 42 (1) (1993) 471.
2 Hong S Y & Zaho, *Cln Prod Prcs*, 1 (1999) 107.
3 Reddy M M &Yi Q S, *J Clean Prod*, 83 (2014) 33.
4 Debnath S, Hong S Y & Broomer M, *Clean Technol Environ Policy*, 2 (2000) 157.
5 Pereira O, Rodriguez A, Fernandez-A A I, Barreiro & Lopezde L L N, *J Clean Prod*, 139 (2016) 440.
6 Shokran A, Dhokia V & Newman S T, *Int J Mach*, 57 (2012) 83.
7 Boubekri N & Foster P R, *J Manuf Syst*, 212 (2015) 556.
8 Tirelli, Chiappini E & Santro M, *Key Eng Mater*, 651 (2014) 1204.
9 Wang Z Y & Raiukra K P, *Wear*, 239 (2000) 168.
10 Gupta M K, Singh G & Sood P K, *J IEI Se C* (2015) https://doi.org/10.007/s40032-015-0175-z.
11 Shokrani A, Dhokia V & Newman S T, *Procedia CIRP*, 46 (2016) 404.
12 Sartori S, Ghiotti A & Brushi S, *Wear*, 376 (2017) 107.
13 Dhananchezian, Rishabapiyan M, Rajashekar G & Narayanan G S, *Mater Today*, 5 (2018) 12062.
14 Sharma A, Singh R C & Singari R M, *Mater Res Express*, 6 (2019) 1.
15 Mia M, Dey P R, Hossain M S, Arafat M T, Asaduzzaman M, Ullah M S & Zobaer S M T, *Meas*, 122 (2018) 380.

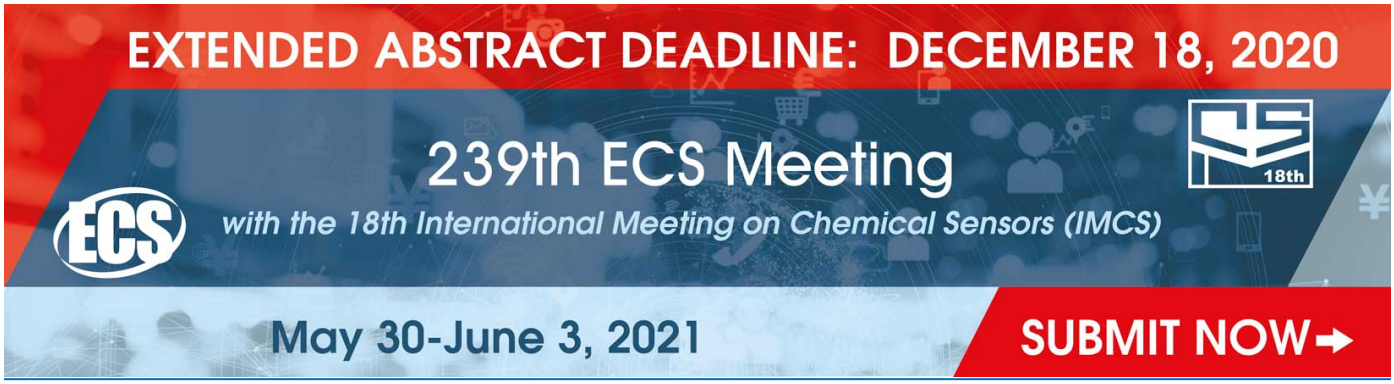
16 Aydin A, Ketenoglu B & Bestanci E, *Indian J Pure Appl Phys*, 58 (2020) 635.
17 Anmol & Pandit S, *Indian J Pure Appl Phys*, 58 (2020) 698.
18 Gupta M K & Sood P K, *J IEI Se C*, 97 (2016) 63.
19 M Vasu & Nayaka H S, *Mater Res Express*, 5 (2018) 501 https://doi.org/10.1088/2053-1591/aac67f.
20 Dr. Vijay K M, Kiran B J & Rudresha N, *Mater Today*, 5 (2018) 11395.
21 Nilrundra M, Doloi B & Mondal B, *J IEI Sec*, 97 (2016) 77.
22 Mia M & Dhar N R, *P I Mech Eng B J Eng* (2017) 1 https://doi.org/10.1177/0954405417737581.
23 Ross P J, *Taguchi Techniques for Quality Engineering, Loss Function, Orthogonal Experiments*, Parameter and Tolerance Design, Newyork McGraw-Hill, 2 (1996) 181.

PAPER • **OPEN ACCESS**

Green Synthesis and Characterization of Copper Oxide Nanoparticles Using *Psidium guajava* Leaf Extract

To cite this article: Ammu Varughese *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **961** 012011

View the article online for updates and enhancements.



This content was downloaded from IP address 14.139.251.100 on 04/12/2020 at 05:53

IOP Conf. Series: Materials Science and Engineering **961** (2020) 012011 doi:10.1088/1757-899X/961/1/012011

Green Synthesis and Characterization of Copper Oxide Nanoparticles Using *Psidium guajava* Leaf Extract

Ammu Varughese[1], Raminder Kaur[1] & Poonam Singh[1]

[1]Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India

ammuvarughese319@gmail.com, raminderkaur@dce.ac.in, poonam@dtu.ac.in

Abstract Green synthesis of nanoparticles has gained enormous attention in today's world due to ongoing demand to develop safe, sustainable, cost-effective and environmentally friendly process for synthesizing a wide variety of materials including metal/metal oxides nanoparticles, hybrid and bioinspired materials. In the current study, we have carried out green synthesis of copper oxide (CuO) nanoparticles (NPs) using *psidium guajava* leaf extract as capping agent and copper acetate as metal precursor. The biosynthesized nanoparticles were characterized using wide variety of techniques i.e. powder X-ray diffraction (PXRD), FTIR, TGA and HRTEM. The PXRD result confirms the synthesis of copper oxide nanoparticles in pure phase having monoclinic symmetry. The average particle size using PXRD was found out to be ~33 nm.

Key words: *ecofriendly, cost-effective, sustainable, nanoparticles and metal oxides*

1. Introduction

Over the last few decades, synthesis of nanomaterials (such as metal NPs, carbon nanotubes (CNTs), quantum dots (QDs), graphene, etc.) having dimension within the range of 10-100

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

nm have become an extensive area of research due to their potential applications in broad area of science and technology [1-6]. Though bulk synthesis of nanoparticles can be carried out rapidly using conventional methods (i.e. physical and chemical methods), they require use of hazardous chemicals [7-9]. Therefore keeping this in mind, use of green synthesis approach that makes use of mild reaction conditions and non-hazardous precursors have been emphasized for promoting environmental sustainability [10-15]. The biological systems that are actively used in the green synthesis of NPs include plants and their derivatives (phytosynthesis), microorganisms like fungi, bacteria, algae and yeast [16-17]. Among all the available methods, phytosynthesis is the ideal method for the synthesis of metal oxide NPs owing to several benefits associated with it i.e. safe, efficient, cost-effective, simple (does not involve complex & multistage process like microbial isolation, culturing, etc.) and provides better control over shape, size & dispersion of metal nanoparticles in comparison to bacteria and/or fungi mediated synthesis [18].

Among a variety of transition metal oxides, CuO is one amongst the potential *p*-type semiconductors that have gained considerable attention attributed to its excellent physical, electrical, optical, and magnetic properties. CuO with a narrow band gap of 1.2 eV is widely used in catalysis [19-22], electrochemistry [23-24], solar energy conversion [25], sensors/biosensors [26-28], energy storage [29], biocidal agents [30-31] etc. In literature, green synthesis of CuO NPs is reported using plants such as *Calotropis procera* leaf extract [32], *Abutilon indicum* leaf extract [33], aloe vera leaf extract [34], karaya gum [35], Cassia alata flower extract etc.

In the current study, *Psidium guajava* leaves have been used for carrying out the synthesis of copper oxide NPs due to the presence of high amount of phytochemicals in it [36-37]. *Psidium guajava* is a medicinal plant commonly known as guava & has a place in Myrtle family as Myrtaceae [38]. *Psidium guajava* is found in various areas of the tropical and subtropical zones and has been reported to have antidiarrheal [39], antispasmodic [40], antioxidant [41], anti-allergy [42], antibacterial [43], anti-cough [44], anti-inflammatory [45], and anticancer activities [46].

2. Materials and methods

Copper acetate monohydrate (Merck, ≥ 98.0 %), Sodium lauryl sulfate (Sisco Research Laboratories ≥ 98.0 %) and fresh *psidium guajava* leaves collected from the Delhi Technological University (DTU) campus.

2.1. Extraction of psidium guajava Leaf Extract:

Fresh *psidium guajava* leaves collected from the DTU campus were first cut into small pieces. These leaves were then washed thoroughly using double distilled water to eliminate debris. These cleaned leaves were then dried in the presence of sunlight for 5 days. After that the dried leaves were ground into a fine powder using pestle mortar & stored in dry place.

Psidium guajava leaf extract was prepared by heating the solution of *psidium guajava* leaf powder and distilled water (1:10 wt. %) at 100°C for 30 minutes in Erlenmeyer flask till the colour of solution changes to dark brown (indicating phytochemical extraction). After that the solution was allowed to cool down to room temperature and filtered through Whatman No.1 filter paper. The filtrate was used for the synthesis of copper oxide NPs.

2.2. Synthesis of CuO Nanoparticles:

To synthesise CuO NPs 1 M copper acetate monohydrate was dissolved in 10 ml of distilled water at room temperature using magnetic stirrer (250 rpm). After that, 30 ml of *psidium guajava* leaf extract was added into the copper acetate solution drop wise, with continuous stirring. Later the solution was kept on magnetic stirrer (250 rpm) for 24 hrs at 60 °C leading to the formation of gel that was dried in hot air oven at temperature of 60 °C. The green precipitate thus obtained was calcined at temperatures of 400 °C for 6 hrs in the muffle furnace, forming black CuO nanoparticles.

2.3. Synthesis of CuO Nanoparticles in presence of sodium lauryl sulfate (SLS)

For the synthesis of CuO nanoparticles in presence of SLS, SLS (4 g) was added into p*sidium guajava* leaf extract (30 ml) and the resulting solution was stirred for five minutes. Rest of the procedure followed was same as mentioned above.

PXRD pattern of synthesized sample was recorded using high-resolution Bruker D8 Advanced X Ray diffractrometer employing Cu Kα radiation (λ = 1.5418 Å) over the range of 2θ = 10-80°. Thermo gravimetric analysis (TGA) was conducted using the Pyris 1® TGA in the range 50-800°C under flowing nitrogen (50 mL min^{-1}) at a heating rate of 10 °C min^{-1}. FTIR spectra of the sample was recorded using a Perkin-Elmer 2000 FTIR spectrometer

employing KBr disks. TEM of the sample was recorded using Jeol/JEM 2100 HRTEM operated at an acceleration voltage of 200KV.

3. Results and discussion

3.1. Qualitative Phytochemical Screening:

Phytochemical screening of freshly prepared *psidium guajava* leaf extract was carried out using simple chemical tests to identify the presence of active phytoconstituent i.e. polyphenol, alkanoids, flavonoids, saponins, tannins, etc. in the sample (Table 1).

Table 1 Phytochemical screening of *psidium guajava* leaf extract

Component	**Procedure**	**Result**
Flavonoids	A portion of aqueous extract was added to 5 ml of aq ammonia solution, followed by the addition of Conc. H_2SO_4 [47].	Absence of yellow colour indicating the absence of flavonoids.
Gum & Mucilages	10 ml of extract was added to 25 ml of ethanol with constant stirring [48].	No white or cloudy precipitate indicating the absence of Gum & Mucilages
Phenolic Compounds	To 5 ml of leaf extract few drops of 5% neutral $FeCl_3$ was added [49].	formation of dark green colour indicating the presence of Phenolic Compounds
Terpenoids	To 5 ml of leaf extract 2 ml of chloroform was added followed by the addition of 2 ml of Conc. H_2SO_4 carefully along the sides of the test tube to form a layer [50].	No formation of a reddish brown colour at the interface indicates the absence of terpenoids.
Saponins	A 1 ml aliquot of the leaf extract was diluted using 20 ml of distilled water and was then shaken vigorously for 15 minutes in a vortex [49].	Persistent foaming indicates the presence of saponins.
Glycosides	Few drops of aqueous NaOH was added into a 1ml of leaf extract [47].	No formation of yellow color indicates the absence of Glycosides.

Steroids	5 ml of extract was added to a 2 ml of chloroform & Conc. H_2SO_4 solution and was shaken thoroughly [51].	No formation of a reddish brown colour at the lower chloroform layer indicates the absence of steroids.

Phytoconstituent are involved in the green synthesis of metal oxide nanoparticles and can acts as both effective reducing as well as stabilizing agents. Therefore in order to understand the mechanism of nanoparticles synthesis screening of phytochemical is inevitable. Phytochemical analysis exhibit positive result for phenolic & saponins content in the leaf extract inferring its significant properties and shows negative results for flavonoids, gum & mucilages, terpenoids, glycosides and steroids.

3.2. Structural Analysis of CuO Nanoparticles

The PXRD pattern of CuO nanoparticles synthesized using *psidium guajava* leaf extract in absence and in presence of SLS is presented in Figure 1. The presence of PXRD peaks at 2θ value of 32.53°, 35.63°, 38.74°, 48.83°, 53.56°, 58.36°, 61.51°, 65.85°, 66.42°, 68.11°, 72.34° and 75.33° confirms the formation of CuO having monoclinic symmetry with lattice constant a = 4.683, b = 3.428, c = 5.129 Å (JCPDS 80-1268) [52]. Since no additional peak due to the presence of any other phase was observed in absence of SLS this suggests the formation of copper oxide nanoparticles in pure phase.

The average particle size of synthesized copper oxide nanoparticles was calculated using Debye- Scherrer's formula, i.e.,

$$\mathbf{D} = \frac{\mathbf{k\lambda}}{\mathbf{\beta Cos\theta}}$$

Where, D, λ, β, θ represents average particle size (nm), wavelength of x-ray (0.15406 nm), full width half maximum of the intense peak, Bragg's angle and k is taken as 0.89 (constant), respectively. Using the above equation, the average particle size of copper oxide nanoparticles was found out to be ~33 nm.

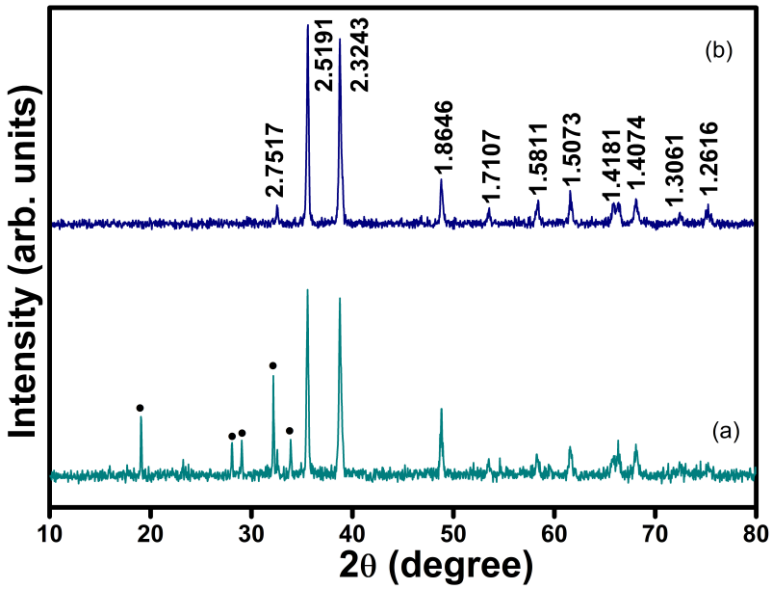


Figure 1: PXRD pattern of CuO nanoparticles in presence (a) and in absence of SLS (b). ''•'' denote reflections due to Na_2SO_4

Further confirmation of internal structure, accurate measurement of particle size and morphology was done using TEM analysis (Figure: 2). Both the average particle size & interplanar spacing values were in close agreement to that obtained from XRD data. However SAED pattern confirms the crystalline nature of synthesized copper oxide nanoparticles.

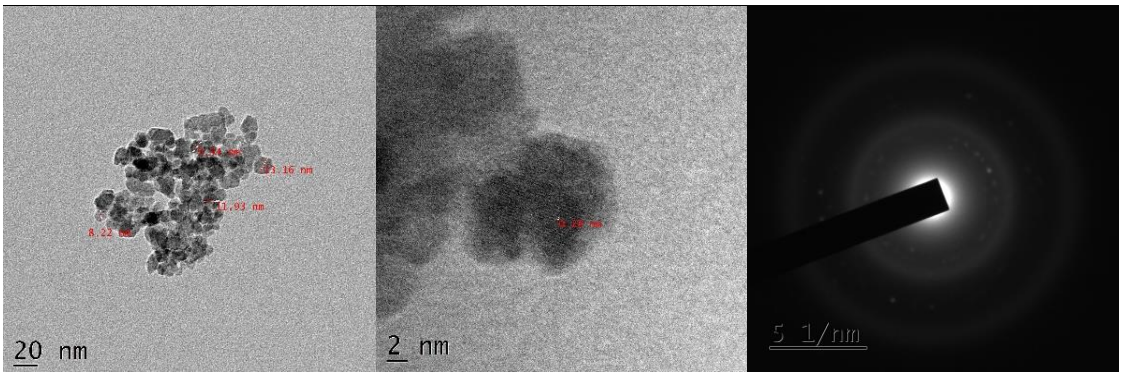


Figure 2: HR-TEM image and SAED pattern of CuO nanoparticles

The FTIR spectrum of synthesized copper oxide nanoparticles is presented in Figure 3. The peaks around 784, 624 and 529 cm^{-1} corresponds to the Cu–O stretching vibration of copper oxide nanoparticles in the monoclinic structure. The absorption peaks at 3439 cm^{-1} and 1628 cm^{-1} corresponds to the OH stretching vibration and HOH bending mode of adsorbed water molecules, since the nano crystalline materials possess high surface area to volume ratio

leading to the absorption of moisture in the lattice. The absorption band at 1107cm^{-1} corresponds to the C–O stretching of phenol and alcoholic compounds.

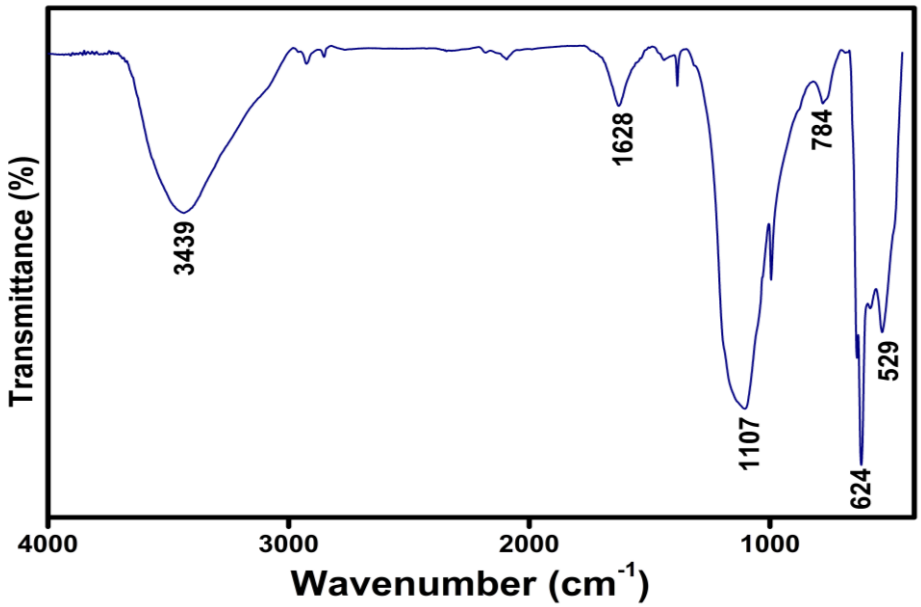


Figure 3: FTIR spectrum of CuO nanoparticles

The TGA/DTG profile of copper oxide nanoparticles is presented in Figure 4. The TGA curve showed a three-step decomposition process of the precursor to form CuO NPs. The first weight loss below 166°C (~9%) corresponds to the removal of moisture and organic solvent molecules. The second and third weight loss in the region from 166 – 403°C may be due to pyrolysis and combustion of organic compounds form precursor and solvent that leaves the system in the form of CO, CO_2 and other organic gases. At further higher temperatures, no significant weight loss was observed, thereby supporting the formation of copper oxide NPs with high purity.

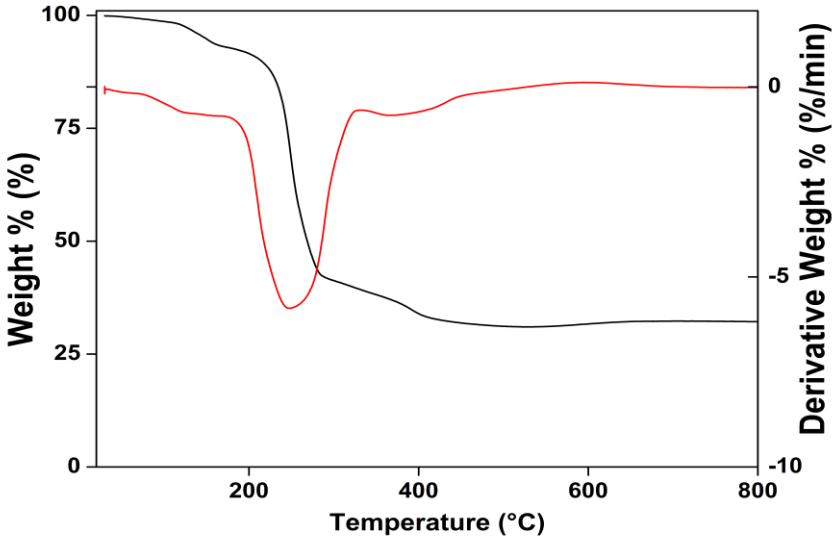
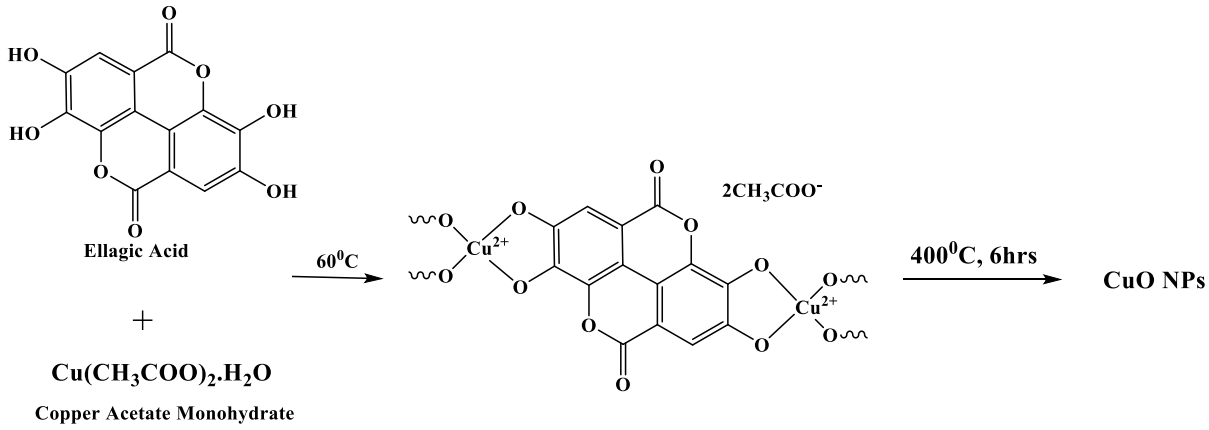


Figure 4: TGA/DTG thermograph of CuO precursor (under N_2 atmosphere)

Hence, based on TGA and phytochemical analysis, the mechanism of biochemical reduction & stabilization of CuO nanoparticles might be as follows:



(**Figure 5:** Mechanism of the synthesis of CuO nanoparticles)

4. Conclusion

The present study reports the biosynthesis of copper oxide nanoparticles by making use of using *psidium guajava* leaf extract as both capping & reducing agent. The presence of phytoconstituent (phenolic compounds) in the leaf extract plays major role in the biochemical reduction & stabilization of copper oxide at nano scale level. The PXRD spectrum confirms the formation of CuO in monoclinic phase having average crystallite size of ~33nm. The HRTEM revealed spherical morphology of CuO nanoparticles. The methodology used for the

synthesis is green and viable because of its ease, efficiency, cost-efficient and eco-friendly nature in comparison to other methods.

5. References

1) Y. Yang, C. Zhang & Z. Hu, **2013**, "Impact of metallic and metal oxide nanoparticles on waste water treatment and anaerobic digestion", *Environmental Science: Processes & Impacts*, **15**, 39.
2) T.W.Woolerton, S.Sheard, E.Pierce, S.W.Ragsdaleb & F.A.Armstrong, 2011, "CO_2 photoreduction at enzyme-modified metal oxide nanoparticles", Energy Environ. Sci., **4**, 2393–2399.
3) S.A. Corr, 2013, "Metal Oxide Nanoparticles", *Nanoscience*, **1**, 180–207.
4) R.Fang, P.Tian, X.Yang, R.Luque, Y.Li, 2012, "Encapsulation of Ultrafine Metal-Oxides Nanoparticles within Mesopores for Biomass-Derived Catalytic Applications", *J. Chem, Science*, **00**, 1-3.
5) Z.Zhang, X.Zou, L.Xu, L.Liao, W.Liu, J.Ho, X.Xiao, C.Jiang, J.Li, 2015, "Hydrogen gas sensor based on metal oxide nanoparticles decorated graphene transistor", *Nanoscale*.
6) N. Verma, N. Kumar, "Synthesis & Biomedical Application of Copper Oxide Nanoparticles: An Expanding Horizon, 2019, *ACS Biomater. Sci. Eng*. **5**, 1170-1188.
7) K.B. Narayanan, N. Sakthivel, "Biological synthesis of metal nanoparticles by microbes", 2010, *Elsevier Advances in Colloid and Interface Science*, **156**, 1–13.
8) P. Dauthal, M. Mukhopadhyay, 2016, "Nobel Metal Nanoparticles: Plant-mediated Synthesis, Mechanistic Aspects of Synthesis, and Applications", *Ind. Eng. Chem. Res*, **55**, 9557-9577.
9) C.P. Devatha, A.K. Thalla, "Green Synthesis of Nanomaterials", *Synthesis of Inorganic Nanomaterials*, 169-184.
10) Anastas P. T, Warner J. C., 1998, Green Chemistry: Theory and Practice; Oxford University Press, Inc.: New York.
11) P. Malik, Ravi Shankar, V. Malik, N Sharma, T. K. Mukherjee, 2014, "Green Chemistry Based Benign Routes for Nanoparticle Synthesis", *Journal of Nanoparticles*, **Volume 2014**, Article ID 302429, 14 pages.

12) J. Virkutyte, R.S. Varma, 2013, "Green Synthesis of Nanomaterials: Environmental Aspects", *ACS Symposisum Series*.
13) P. Raveendran, J. Fu, S.L. Wallen, 2003, "Completely "Green" Synthesis and Stabilization of Metal Nanoparticles", *J. AM. CHEM. SOC*., **125,** 13940-13941.
14) V. Srivastava, D. Gusain, Y.C. Sharma, 2015, "Critical Review on some widely used Engineered Nanoparticles", *Ind. Eng. Chem. Res*, **54**, 6209-6233.
15) D.L. Hjeresen, D.L. Schutt, J.M. Boese, 2000, "Green Chemistry & Approach", *Journal of Chemical Education*, **Vol 77**, No. 12, 1543-1547.
16) S. Pirtarighat, M. Ghannadnia, S. Baghshahi, 2019, "Green synthesis of silver nanoparticles using the plant extract of Salvia spinosa grown in vitro and their antibacterial activity assessment", *Journal of Nanostructure in Chemistry*, **9**, 1–9.
17) P. Khandel, R.K. Yadaw, D.K. Soni, L. Kanwar, S.K. Shahi, 2018, "Biogenesis of metal nanoparticles and their pharmacological applications: present status and application prospects", *Journal of Nanostructure in Chemistry*, **8**, 217–254.
18) M.S. Akhtar, J. Panwar, YS Yun, 2013, "Biogenic Synthesis of Metallic Nanoparticles by Plant Extracts", *ACS Sustainable Chem. Eng.,* **1**, 591-602.
19) Z. Su, J. Li, D. Zhang, P. Ye, H. Li, Y. Yan, 2019, "Novel flexible Fenton-like catalyst: Unique CuO nanowires arrays on copper mesh with high efficiency across a wide pH range", *Science of the Total Environment*, **647**, 587–596.
20) M.A. Bhosale, B.M. Bhanage, 2014, "A facile one-step approach for the synthesis of uniform spherical Cu/Cu_2O nano- and microparticles with high catalytic activity in the Buchwald–Hartwig amination reaction", *RSC Adv*., **4**, 15122.
21) S.Momeni, F.Sedaghati, 2018, "CuO/Cu_2O nanoparticles: A simple and green synthesis, characterization and their electrocatalytic performance towards formaldehyde oxidation", *Microchemical Journal*, **143**, 64.
22) X. Liu, H. Jia, Z. Sun, H. Chen, P. Xu, P. Du, 2014, "Nanostructured copper oxide electrodeposited from copper(II)complexes as an active catalyst for electrocatalytic oxygen evolution reaction", *Electrochemistry Communications,* **46**, 1-4.
23) C.W. Li, J. Ciston, M.W. Kanan, 2014, "Electroreduction of carbon monoxide to liquid fuel on oxide-derived nanocrystalline copper", *Nature*, **Vol 000**, 1-4.

24) A. Pendashteh, M.F. Mousavi, M.S. Rahmanifar, 2013, "Fabrication of anchored copper oxide nanoparticles on graphene oxide nanosheets via an electrostatic coprecipitation and its application as Supercapacitor", *Electrochimica Acta*, **88**, 347– 357.
25) H. Siddiqui, M.R. Parra, P. Pandey, M.S. Quereshi, F. Z. Haque, 2020, "Utility of copper oxide nanoparticles (CuO-NPs) as efficient electron donor material in bulk-heterojunction solar cells with enhanced power conversion efficiency", *Journal of Science: Advanced Materials and Devices*, **Vol. 5**, Issue 1, 104-110.
26) A.D. Tocco, S.N. Robledo, Y. Osuna, J. Sandoval-Cortez, A.M. Granero, N.R. Vettorazzi, J.L. Martínez, E.P. Segura, A. Iliná, M.A. Zon, F.J. Arévaloa,, H. Fernández, 2018, "Development of an electrochemical biosensor for the determination of triglycerides in serum samples based on a lipase/magnetite-chitosan/copper oxide nanoparticles/multiwalled carbon nanotubes/pectin composite", *Talanta*, **190**, 30–37.
27) T. Alizadeh, S. Mirzagholipur, 2014, "A Nafion-free nonenzymatic amperometric glucose sensor based on copper oxide nanoparticles-graphene nanocomposite", *Sensors and Actuators B*, **198**, 438-447.
28) F.E. Annanouch, Z. Haddi, S. Vallejos, P. Umek, P. Guttmann, C. Bittencourt, E. Llobet, 2015, "Aerosol assisted CVD grown WO3 nanoneedles decorated with copper oxide nanoparticles for the selective and humidity resilient detection of H2S", ACS Appl. Mater. Interfaces, **7**, 12, 6842-6851.
29) P. Liang, F. Wang, Z. Hu, 2018, "Controlled synthesis of ordered sandwich CuCo2O4/reduced graphene oxide composites via Layer-by-layer hetero-assembly for high-performance supercapacitors", *Chemical Engineering Journal*, **350**, 627.
30) S.M. Dizaj, F. Lotfipour, M.B. Jalali, M.H. Zarrintan, K. Adibkia, 2014, "Antimicrobial activity of the metals and metal oxide nanoparticles", *Materials Science and Engineering*, **44**, 278–284.
31) R. Sivaraj, P.K.S.M. Rahman, P. Rajiv, H.A. Salam, R. Venkatesh, 2014, "Biogenic copper oxide nanoparticles synthesis using Tabernaemontana divaricate leaf extract and its antibacterial activity against urinary tract pathogen", *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **133**, 178–181.
32) K. Rayapa Reddy, 2017, "Green synthesis, morphological and optical studies of CuO nanoparticles", *Journal of Molecular Structure*, **1150**, 553-557.

33) F. Ijaz, S. Shahid, S.A. Khan, W. Ahmad, S. Zaman, 2017, "Green synthesis of copper oxide nanoparticles using Abutilon indicum leaf extract: Antimicrobial, antioxidant and photocatalytic dye degradation activities", *Tropical Journal of Pharmaceutical Research*, **16**, 4: 743-753.

34) P. P. N. Vijay Kumar, U. Shameem, P. Kollu, R. L. Kalyani & S. V. N. Pammi, 2015, "Green Synthesis of Copper Oxide Nanoparticles Using Aloe vera Leaf Extract and Its Antibacterial Activity against Fish Bacterial Pathogens", *BioNanoScience,* **vol. 5**, 135–139.

35) Vinod Vellora Thekkae Padil & Miroslav Černík, 2013, "Green synthesis of copper oxide nanoparticles using gum karaya as a biotemplate and their antibacterial application", *International Journal of Nanomedicine*, **8**, 889–898.

36) E. D-Cerio, AM G-Caravaca, V Verardo, A F-Gutierrez & A S-Carretero, 2016, "Determination of guava (Psidium guajava L.) leaf phenolic compounds using HPLC-DAD-QTOF-MS", *Journal of Functional Foods*, **22**, 376–388.

37) E. D-Cerio, AM G-Caravaca, V Verardo, A F-Gutierrez & A S-Carretero, 2015, "Determination of Polar Compounds in Guava Leaves Infusions and Ultrasound Aqueous Extract by HPLC-ESI-MS", *Journal of Chemistry*, **Volume 2015**, Article ID 250919, 9 pages.

38) T. Santhoshkumar, AA Rahman, C Jayaseelan, G Rajakumar, S Marimuthu, AV Kirthi, K Velayutham, J Thomas, J Venkatesan, SK Kim, 2014, "Green synthesis of titanium dioxide nanoparticles using Psidium guajava extract and its antibacterial and antioxidant properties", *Asian Pacific Journal of Tropical Medicine*, 968-976.

39) Ojewole JA, Awe EO, Chiwororo WD, 2008, "Antidiarrheal activity of Psidium guajava Linn. (Myrtaceae) leaf aqueous extract in rodents", *J Smooth Muscle Res*, **44**, 6, 195-207.

40) X. Lozoya, H. Reyes-Morales, M.A Chávez-Soto, M.C. Martínez-García, Y. Soto-González, S.V. Doubova, 2002, "Intestinal anti-spasmodic effect of a phytodrug of Psidium guajava folia in the treatment of acute diarrheic disease", *J Ethnopharmacol*, **83**, 1-2, 19-24.

41) S.M. El-Amin, M.A.M. Hashash, A.M. Abdou, A.M. Saad, M.S. Abdel-Aziz, A.S. Mohamed, 2016, "Antimicrobial and antioxidant activities of Psidium guajava leaves growing in Egypt", *Der Pharmacia Lettre*, **8**, 12, 27-33.

42) N. Seo, T. Ito, N. Wang, X. Yao, Y. Tokura, F. Furukawa, M. Takigawa, S Kitanaka, 2005, "Anti-allergic Psidium guajava extracts exert an antitumor effect by inhibition of T regulatory cells and resultant augmentation of Th1 cells", *Anticancer Res*, **25**, 6A, 3763-70.

43) Neviton RS, Aparicio DGC, Simone MS, Vataru CN, Prado BF, 2005, "An evaluation of antibacterial activity of Psidium guajava (L)", *Braz Arch Biol Technol*, **8**, 3, 429-436.

44) P. Jaiarj, P. Khoohaswan, Y. Wongkrajang, P. Peungvicha, P. Suriyawong, M.L. Sumal Saraya, O. Ruangsomboon, 1999, "Anticough and antimicrobial activities of Psidium guajava Linn. leaf extract", *Journal of Ethnopharmacology*, **Vol 67**, 2, 203-212.

45) Ojewolw JA, 2006, "Antiinflammatory and analgesic effects of Psidium guajava Linn (Myrtaceae) leaf aqueous extracts in rats and mice", *Meth Find Exp Clin Pharmacol*, **28** 7, 441-446.

46) Sang-Bang L, Hae-Ryong P, 2010, "Anticancer activity of guava (Psidium guajava L.) branch extract against HT-29 human colon cancer cells", *J Med Plant Res*, **4**, 10, 891-896.

47) Emasushan M, John Britto S, 2018, "Preliminary phytochemical profiling and antifungal activity of the seeds and pericarp of Putranjiva roxburghii Wall", *The Pharma Innovation Journal*, **7**, 4, 107-110.

48) C. Uma, K.G. Sekar, 2014, "Phytochemical analysis of a folklore medicinal plant citrullus colocynthis L (bitter apple)", *Journal of Pharmacognosy and Phytochemistry*, **2**, 6, 195-202.

49) K.S. Banu, L. Catherine, 2015, "General Techniques Involved in Phytochemical Analysis", *IJARCS*, **2**, 4, 25-32.

50) R.S. Kumar, C. Venkateshwar, G. Samuel, S.G. Rao, 2013, "Phytochemical Screening of some compounds from plant leaf extracts of Holoptelea integrifolia (Planch.) & Celestrus emarginata (Grah.) used by Gondu tribes at Adilabad District, Andhrapradesh, India", International Journal of Engineering Science Invention, **Volume 2**, Issue 8, PP.65-70.

51) R. Gul, S.U. Jan, S. Faridullah, S. Sherani, N. Jahan, 2017, "Preliminary Phytochemical Screening, Quantitative Analysis of Alkaloids, and Antioxidant Activity of Crude Plant Extracts from Ephedra intermedia Indigenous to Balochistan", *Hindawi Scientific World Journal*, **Volume 2017**.

52) J. Zhu, D. Li, H. Chen, X. Yang, L. Lu, X. Wang, 2004, "Highly dispersed CuO nanoparticles prepared by a novel quick-precipitation method", *Materials Letters*, **58**, 3324 – 3327.

Acknowledgments

I acknowledge my deep sense of gratitude to my research advisors Dr. Raminder Kaur & Dr. Poonam Singh, Assistant Professor, Department of Applied Chemistry, Delhi Technological University, Delhi for their guidance, motivation and remarkable patience throughout this research work. Also, I am grateful to the research scholars, Ms. Bhamini Pandey & Ms. Jigyasa Pathak, and the technical staffs of Department of Applied Chemistry, Delhi Technological University, Delhi for their constant support and assistance during this research work.

Soft Computing (2020) 24:18287–18300
https://doi.org/10.1007/s00500-020-05367-9

FOUNDATIONS

Group operations and isomorphic relation with the 2-tuple linguistic variables

Tanya Malhotra[1] · **Anjana Gupta**[1]

Published online: 5 November 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
This paper aims to put forth the theory of 2-tuple linguistic groups concerning the binary operation in the conventional sense. For this, a formal methodology has been introduced to prove that a predefined nonempty linguistic term set, LT, and the interval, $[\frac{-1}{2}, \frac{1}{2}]$, forms a group. Further, we have proved that a set of all 2-tuple linguistic information, $\overline{LT} \equiv LT \times [\frac{-1}{2}, \frac{1}{2}]$, and numerical interval, $[-n, n]$, where n is presumed to be a positive integer, also forms a group. Later on, we develop a one-to-one correspondence and homomorphic group relation between the set of all 2-tuple linguistic information and numerical interval, $[-n, n]$. Henceforth, a similarity relation between the two groups is obtained. Finally, a practical application is defined by proposing the notion of a 2-tuple linguistic bipolar graph to illustrate the usefulness and practicality of the group isomorphic relation.

Keywords Subscript-symmetric linguistic model · Linguistic group operation · 2-tuple linguistic group isomorphism · 2-tuple linguistic model · 2-tuple linguistic bipolar set · 2-tuple linguistic bipolar graphs

1 Introduction

The Zadeh's (1965) introduction of fuzzy sets is considered to be the most relevant concept in describing the problems of real-life since every object in the physical world possesses some degree of fuzziness. Therefore, fuzzy-based approaches are well-suited to deal with uncertainty and have become a considerable research area with manifold applications in respect of engineering, computer science, and so forth. Several investigations have also been envisioned to "fuzzify" many essential mathematical structures like algebra, topological spaces, groups, etc.

It is a well-known fact that the classical theory of groups is recognized as the study of algebraic structures designed to model and thus contemplate the symmetries of specific objects. In other words, group theory plays a notable role in mathematics to analyze individual objects' symmetries. An object's symmetry can be either discrete or continuous. It is, therefore, imperative to examine each discrete and continuous group. Ideally, group theory evolution begins in 1770 and is extended to the twentieth century. However, the nineteenth century witnessed the significant era of development of group theory. Since then, the scope of group theory has been expanded to geometry, cryptography, particle physics, combinatorics, some areas of analysis, and so forth. Numerous researchers have contributed considerably in order to develop theories and outlooks to enrich the wide-ranging application of group theory based on a crisp set (to mention only a few, see Kleiner 1986; Kahrobaei and Anshel 2010; Myasnikov et al. 2008; Rodríguez-Lara et al. 2018; Dietrich and Wilson 2019).

The scope of group theory is further anticipated in terms of fuzzy variables and fuzzy sets (Zadeh 1965) widely referred to as the fuzzy group theory. Rosenfeld (1971) initially introduced the idea of fuzzy groups to outspread the elementary concepts of classical group theory in an effort to reinforce fuzzy groups. Subsequently, (De Luca and Termini 1972) defined appropriate algebraic properties associated with the class of fuzzy sets as well as making a careful examination of certain new algebraic aspects of the "fuzzy sets" theory, thereby linking it with Brouwerian lattices. Furthermore, (Anthony and Sherwood 1979) redefined the concept of fuzzy algebraic structures based on the kind of semi-groups defined on the unit interval [0, 1] called as "t-norm" introduced by

Communicated by A. Di Nola.

✉ Anjana Gupta
anjanagupta@dce.ac.in

[1] Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India

Schweizer and Sklar (1960). In that paper, authors have stated new results and properties in light of the proposed definition and, consequently, overcomes the limitation possessed by the previously stated fuzzy algebraic concept (Rosenfeld 1971; De Luca and Termini 1972).

Das (1981) inspected fuzzy groups to characterize all the fuzzy subgroups of finite cyclic groups and put forward the idea of "level subgroups" of a fuzzy subgroup centered on a concept of "level subset" developed by Zadeh. For a more detailed knowledge about the results and deeper understanding of a concept, a reader can refer to the corresponding paper (Das 1981). Mukherjee and Bhattacharya (1984) gave an introduction of a "fuzzy normal subgroup" and "fuzzy coset" and have further established a fuzzy Lagrange's theorem. Thereby, presented the analogs of the basic group-theoretic results in an elementary fashion. In a sequel, Mukherjee and Bhattacharya (1986); Bhattacharya and Mukherjee (1987) further investigated several analogous concepts of classical group theory in terms of fuzzy groups and proved the fuzzy variants of a few prominent group theoretical results. Also, (Bhattacharya 1987) proposed a more detailed study about the "level subgroups" of a fuzzy subgroup and further examined the possibility of the unique determination of the fuzzy subgroup by the class of "level subgroups" of a fuzzy subgroup. The author also rectified the inaccuracies persisting within the few results of Das (1981) and later on initiated to delineate all the fuzzy subgroups of a group of finite order in an effort to generalize the results of Rosenfeld (1971) and Das (1981).

Sherwood (1983) presented the concept of product of fuzzy groups and further discussed the related properties as well. Later on, (Ray 1992) gave the introduction of the isomorphism of fuzzy groups and have stated an acceptable definition of the fuzzy version of the isomorphism theorems. Analogous to classical normal groups and group homomorphism, (Akgül 1988) defined "fuzzy normal subgroups," "fuzzy-level normal subgroups," and their homomorphism. In a sequel, (Makamba 1992) introduced the notion of the internal and external direct product of a fuzzy subgroup of a corresponding group G and has additionally established fuzzy isomorphic relation between fuzzy subgroups. However, apart from the literature mentioned above, readers can also refer to the papers (to mention a few, see Tărnăuceanu and Bentea 2008; Degang et al. 2005; Demirci and Recasens 2004; Yuan and Lee 2004; Ray 1999; Gupta and Sarma 1999; Kim 1997; Asaad and Abou-Zaid 1996; Ovchinnikov 1995; Tu and Wenxiang 1994) to understand in more detail about concepts of fuzzy algebra.

Although "fuzzy sets" theory and fuzzy algebra have been massively studied and become a robust field of research in severe engineering branches, medical sciences, graph theory, etc., there are persisting some real-life situations where "fuzzy sets" theory is not suitable to handle uncertainty. This happens due to the presence of much enigmatic and imprecise information in intricate problems. As a result, it is quite clumsy and unnatural to express results quantitatively in terms of a fuzzy variable. Therefore, the use of linguistic variable and linguistic-based approaches, as developed by Zadeh (1975a, b, c), is much suitable and straightforward to handle uncertainty and thus has been studied extensively (see Zadeh 1975a, b, c, 1976; Tong and Bonissone 1980; Delgado et al. 1998, 1992; Bordogna et al. 1997; Herrera et al. 1995, 1996). Several decision-making problems are solved using linguistic information to model experts' preferences, which implies using computing with words (CWWs) processes. To efficiently address linguistic information, different approaches have been put forward in the literature, namely: (1) linguistic computational models based on membership function (Degani and Bortolan 1988; Dubois and Prade 1980; Klir and Yuan 1995; Bonissone and Decker 1986) (2) ordinal scales approach (Yager 1981). Because of the retranslation step performed by these so-called models to manifest the outcomes in the expression domain originally defined, as a consequence, it leads to an information loss and thus causes insufficiency of accuracy within the final result. However, to overcome the limitation possessed by the previously stated models (Herrera and Martínez 2000) proposed a pioneering linguistic computational model, namely "2-tuple linguistic model" so as to further authenticate the linguistic computations.

Since in meaning, a linguistic variable is in sight as a fuzzy set and 2-tuple representation is alleged to be the most effective way of representing linguistic information. Therefore, in the present work, we link the theory of 2-tuple linguistic set and classical group theory to develop the notion of 2-tuple linguistic groups under binary operations in a classical sense. Consequently, this paper can be considered as an extension of the fuzzy algebra, particularly fuzzy groups. For this, we have developed a methodology in proving that a finite subscript-symmetric linguistic term set (SSLTS), intervals $[-n, n]$ and $[\frac{-1}{2}, \frac{1}{2}]$ forms a group under a binary operator '$*$', '$\oplus$', respectively. In addition, we have proved that $\overline{LT} \equiv LT \times [\frac{-1}{2}, \frac{1}{2}]$ is a 2-tuple linguistic direct group under a binary operation '$\circ$'. The definition of all these group operations are defined in the subsequent section.

Lastly, we present a result describing an isomorphic relation between the 2-tuple linguistic group, $\overline{LT}$, and a numerical interval, $[-n, n]$ based on the classical definition of isomorphic groups. To validate the results and showcase the physical meaning of the isomorphic groups, we herein propose an application in terms of bipolar graphs. In this connection, we first developed the concept of 2-tuple linguistic bipolar set (2TLBS) accompanied by the introduction of the notion of 2-tuple linguistic bipolar graphs (2TLBGs) based on a finite SSLTS and after that define its relevant properties. Later on, we define graph isomorphism viewed as a direct

consequence of a 2-tuple linguistic group isomorphic relation.

The work presented in this contribution is advantageous in developing the basis for linguistic group theory. Although the fuzzy group theory has been developed and widely studied in the literature, linguistic group theory development is pristine and yet to be explored. Therefore, the ideas and results exhibited in this paper are a milestone in this direction and also provide a more flexible framework in proving an "algebraic structure groups" for the 2-tuple linguistic model by using SSLTS. Moreover, the paper's results depict an analogy with an existing theory of groups based on a crisp set.

The paper's layout is modified subsequently. In Sect. 2, we provide a brief review of the basic concepts related to the traditional 2-tuple linguistic model, subscript symmetric linguistic computational model and classical group theory. In Sect. 3, we present the necessary results needed to establish the isomorphic relation between the linguistic groups. In Sect. 4, we give an application of the group isomorphic relation previously stated. For this, we first introduce the notion of 2-tuple linguistic bipolar set, followed by the concept of 2-tuple linguistic bipolar graphs. Then, finally, 2-tuple linguistic bipolar graph isomorphism is presented to explain the physical meaning of group isomorphism. Lastly, Sect. 5 provides the concluding remarks of the paper.

2 Preliminaries

The present section aims to put forward some basic definitions and operational laws related to the 2-tuple linguistic model, subscript symmetric linguistic computational model and classical group theory.

2.1 2-tuple linguistic model

Herrera and Martínez' (2000) 2-tuple model effectively describes the linguistic information present within the problem and further ameliorate the linguistic computational techniques. According to this model, linguistic information is represented in the form of paired values popularly referred as "2-tuple". This so-called 2-tuple linguistic label is composed of a linguistic term and a numerical value, i.e., $(t_p, \alpha) \in S \times [\frac{-1}{2}, \frac{1}{2})$, where $t_p \in S, \alpha \in [\frac{-1}{2}, \frac{1}{2})$.

Definition 1 Herrera and Martínez (2000) Let $S = \{t_p \mid p = 0, \ldots, n\}$ be the collection of linguistic terms such that $|S| = n + 1$ and let $x \in [0, n]$. Let $p = \text{round}(x)$, $p \in \{0, \ldots, n\}$ and $\alpha = x - p \in [\frac{-1}{2}, \frac{1}{2})$ be two numeric values, then we call α as "*symbolic translation*," i.e., a numeric number portraying the value of translation from the original value x to the immediate index label 'p'.

This extensively known 2-tuple model is outlined via transformation function defined between numerical values and the 2-tuple previously stated to conveniently take forward the processes of linguistic computations without any information loss.

Definition 2 Herrera and Martínez (2000) Assume $S = \{t_p \mid p = 0, \ldots, n\}$ as a finite set of linguistic terms and let $x \in [0, n]$ be a numeric value, then the conventional 2-tuple representing the identical information to x is defined underneath:

$$\Delta : [0, n] \rightarrow S \times [\frac{-1}{2}, \frac{1}{2})$$
$$\Delta(x) = (t_p, \alpha), \text{ where } p = \text{round}(x) \in S \,\&\, \alpha = x - p \in [\frac{-1}{2}, \frac{1}{2}).$$

where round(*) is considered in general rounding operation.

Remark 1 Herrera and Martínez (2000) The Δ mapping is both one-one and onto and hence invertible, therefore there exist an inverse transformation, i.e., $\Delta^{-1} : S \times [\frac{-1}{2}, \frac{1}{2}) \rightarrow [0, n]$ that effectively facilitate in returning an equivalent value $x \in [0, n] \subset \mathbb{R}$ corresponding to identical 2-tuple such that $\Delta^{-1}(t_p, \alpha) = x = p + \alpha$.

Remark 2 Herrera and Martínez (2000) A linguistic term t_p is easily possible to be converted into a corresponding 2-tuple linguistic term by simply making an addition of a value 0 as "symbolic translation," i.e., $t_p \in S \Rightarrow (t_p, 0) \in S \times [\frac{-1}{2}, \frac{1}{2})$.

Definition 3 Herrera and Martínez (2000) Let $S = \{t_p \mid p = 0, \ldots, n\}$ has a usual meaning, then the negation of a 2-tuple linguistic label is identified as:

$$\text{neg}((t_p, \alpha)) = \Delta(n - (\Delta^{-1}(t_p, \alpha))).$$

2.2 Subscript-symmetric linguistic computational model

Definition 4 Xu (2005) Assume $LT = \{t_{-n}, \ldots, t_0, \ldots, t_n\}$ be a finite subscript symmetric linguistic term set such that the following properties are satisfied:

(i) The set LT is termed as an ordered set, i.e., $t_p > t_q$ if and only if $p > q$;
(ii) The negation operator of a linguistic variable is defined as: $\text{neg}(t_p) = t_{-p} \in LT$.

Here, t_0 represents a middle linguistic term, and the rest over terms are placed uniformly on each side of the middle term. In particular, t_{-n} and t_n represent the lower and upper bound of the linguistic term set.

The existing 2-tuple linguistic model explained briefly in the previous subsection is defined over the domain $[0, n]$. However, this model can be easily specified for the domain $[-n, n]$ by using the above-mentioned SSLTS, namely LT (see Malhotra et al. 2019). Meanwhile, the notion of Δ and Δ^{-1} operators, ideally formalized by Herrera and Martínez (2000), is extended to the set LT and the transformation function facilitating to convert a numeric value x to a corresponding 2-tuple linguistic variable is given below.

Definition 5 Malhotra et al. (2019) Let $LT = \{t_p \mid p = -n \ldots, 0, 1, \ldots, n\}$ be a set of finite linguistic terms having cardinality $2n + 1$ and let $x \in [-n, n]$. Then, the function Δ needed to express the 2-tuple linguistic variable identical to x is identified as.

$$\Delta : [-n, n] \to LT \times \left[\frac{-1}{2}, \frac{1}{2}\right)$$
$$\Delta(x) = (t_p, \alpha) \text{ where } p = \text{round}(x) \in LT \,\&\, \alpha = x - p \in \left[\frac{-1}{2}, \frac{1}{2}\right).$$

where round(*) is termed as the general round operation, t_p represents the linguistic label nearest to x, and α referred as the "symbolic translation".

From the above transformation function, it has been clearly envisioned that the Δ is a bijective mapping (Herrera and Martínez 2000), and therefore, its inverse is stipulated by,

$$\Delta^{-1} : LT \times \left[\frac{-1}{2}, \frac{1}{2}\right) \to [-n, n] \text{ as } \Delta^{-1}(t_p, \alpha) = p + \alpha = x.$$

Furthermore, (Herrera and Martínez 2000) has established the comparison law for the well-known linguistic model of 2-tuple primarily based upon the classical lexicographic ordering. Hence, correspondingly the following definition can be outlined to state the ranking order.

Definition 6 (Malhotra et al. 2019) Let (t_p, α_p) and (t_q, α_q) be two 2-tuple linguistic variables based on the term set LT of finite order. Then,

(i) If $p < q$ then $(t_p, \alpha_p) < (t_q, \alpha_q)$.
(ii) If $p = q$, i.e., $t_p = t_q$, then
  (a) if $\alpha_p = \alpha_q$, then $(t_p, \alpha_p) = (t_q, \alpha_q)$;
  (b) if $\alpha_p > \alpha_q$, then $(t_p, \alpha_p) > (t_q, \alpha_q)$;
  (c) if $\alpha_p < \alpha_q$, then $(t_p, \alpha_p) < (t_q, \alpha_q)$.

For a comprehensive and complete elucidation of the theory for the existing 2-tuple model, the reader is advised to refer the paper (Herrera and Martínez 2000, 2001).

2.3 Classical group theory

In the following, we review the formal definition of the classical group G and some elementary concepts required for this paper.

Definition 7 Gallian (2012) A group G is defined as a nonempty set along with a group operator '$\bullet \in \{+, \cdot\}$' such that it should satisfy the subsequent characteristics:

(i) $g \bullet h \in G \,\forall\, g, h \in G$;
(ii) Associative property holds trivially;
(iii) There exist an element $e \in G$ such that $g \bullet e = e \bullet g = g \,\forall\, g \in G$;
(iv) For every $g \in G$, there exist an inverse element $f \in G$ such that $g \bullet f = f \bullet g = e$.

It can be noted that the inverse and identity element of a group G is uniquely defined.

Note 1 If a group G satisfies an additional condition

$$g \bullet h = h \bullet g$$

for all $g, h \in$ G then G is referred as abelian group.

In a generic way, if $\bullet = +$ then we write $g \bullet h$ as $g + h \in G \,\forall\, g, h \in G$ and the identity element is 0 while the inverse element of g is $-g$. On the other hand if $\bullet = \cdot$ then we write $g \bullet h$ as $g \cdot h \in G \,\forall\, g, h \in G$ and the identity element is 1 while the inverse element of g is g^{-1}.

Note 2 If a group G consists of a finite range of elements, then G is referred to as a finite group; otherwise, we call G an infinite group.

Note 3 The total number of elements present within the group G is referred as order of the group and is denoted by $|G|$.

Definition 8 Gallian (2012) A subset H of G forms a subgroup of G if it satisfies all the group properties under the same binary operation of G and it is denoted by $H \leq G$.

Definition 9 Gallian (2012) A group homomorphism is defined as a mapping $\phi : G \to \overline{G}$ that preserves the operation of the group; i.e., $\phi(g \bullet h) = \phi(g) \# \phi(h) \,\forall\, g, h \in G$ where $\bullet, \# \in \{+, \cdot\}$ are binary operations of group G and $\overline{G}$, respectively.

Definition 10 Gallian (2012) Group isomorphism is defined as a bijective homomorphic mapping ϕ from a group G to group $\overline{G}$, i.e., $\phi : G \to \overline{G}$ such that $\phi(g \bullet h) = \phi(g) \# \phi(h) \,\forall\, g, h \in G$ where $\bullet, \# \in \{+, \cdot\}$ are binary operations of group G and $\overline{G}$, respectively. When group G is isomorphic to $\overline{G}$, then we say that $G \cong \overline{G}$, i.e., both groups are equivalent and seems out to be abstractly similar in several aspects.

Definition 11 Gallian (2012) Let $G_1, G_2, \ldots, G_n$ be a finite collection of arbitrary groups. The direct product of groups $G_1, G_2, \ldots, G_n$ is defined as a set of n-tuples such that ith component is an element of G_i and the operation is component wise.

Mathematically, direct product of groups is written as:

$$G_1 \times G_2 \times \ldots \times G_n = \{(g_1, g_2, \ldots, g_n) \mid g_i \in G_i\},$$

where

$$(g_1, g_2, \ldots, g_n) \bullet (g_1', g_2', \ldots, g_n') = (g_1 \bullet g_1', g_2 \bullet g_2', \ldots, g_n \bullet g_n')$$

and '$\bullet$' is the respective group operations of each G_i .

For a complete knowledge about classical group theory, it is advised to the reader to refer any familiar text on abstract algebra, for instance, (Gallian 2012).

3 Isomorphic 2-tuple linguistic groups

In the present section, we give a comprehensive study to introduce the concept of 2-tuple linguistic groups based upon the binary operator in a crisp sense. To do this, we initially show that $[-n, n]$, LT, the interval $[\frac{-1}{2}, \frac{1}{2}]$ forms a group under the binary operation '$*$', '$\oplus$', respectively, and further, prove the assertion that the set of all 2-tuple linguistic variables, LT $\times[\frac{-1}{2}, \frac{1}{2}]$ is isomorphic to the nonempty set of numerical interval $[-n, n]$, where n is a positive integer. Hence, both the groups are abstractly similar and have identical algebraic properties.

Proposition 1 *Let $[-n, n]$ be any arbitrary nonempty set where n is presumed to be a sufficiently large positive integer. Then $[-n, n]$ is a group concerning binary operation '$*$' defined as:*

$$a * b = \begin{cases} a + b + n, & \text{if } a + b < -n \\ a + b, & \text{if } -n \le a + b \le n \\ a + b - n, & \text{if } a + b > n \end{cases}$$

for all $a, b \in [-n, n]$.

Proof Let $a_p, a_q \in X$ be any arbitrary element. Then,

$$a_p * a_q = \begin{cases} a_p + a_q + n, & \text{if } a_p + a_q < -n \\ a_p + a_q, & \text{if } -n \le a_p + a_q \le n \\ a_p + a_q - n, & \text{if } a_p + a_q > n \end{cases}$$

clearly, $a_p, a_q \in X$. Therefore, X satisfies the closure property. Now, let $a_p, a_q, a_r \in X$ and let $a_q * a_r = \mu_1$ such that $a_p * (a_q * a_r) = a_p * \mu_1 = \rho_1$ therefore,

$$\rho_1 = \begin{cases} a_p + \mu_1 + n, & \text{if } a_p + \mu_1 < -n \\ a_p + \mu_1, & \text{if } -n \le a_p + \mu_1 \le n \\ a_p + \mu_1 - n, & \text{if } a_p + \mu_1 > n \end{cases} \quad (1)$$

since $\mu_1 = a_q * a_r$, i.e.,

$$\mu_1 = \begin{cases} a_q + a_r + n, & \text{if } a_q + a_r < -n \\ a_q + a_r, & \text{if } -n \le a_q + a_r \le n \\ a_q + a_r - n, & \text{if } a_q + a_r > n \end{cases} \quad (2)$$

By substituting Eq. 2 in Eq. 1, we obtain following cases defined below:

(Case(i)) If $a_q + a_r < -n$ then, $\mu_1 = a_q + a_r + n$ therefore,

$$\rho_1 = \begin{cases} a_p + a_q + a_r + 2n, & \text{if } a_p + a_q + a_r < -2n \\ a_p + a_q + a_r + n, & \text{if } -2n \le a_p + a_q + a_r \le 0 \\ a_p + a_q + a_r, & \text{if } a_p + a_q + a_r > 0 \end{cases}$$

(Case(ii)) If $-n \le a_q + a_r \le n$ then, $\mu_1 = a_q + a_r$ therefore,

$$\rho_1 = \begin{cases} a_p + a_q + a_r + n, & \text{if } a_p + a_q + a_r < -n \\ a_p + a_q + a_r, & \text{if } -n \le a_p + a_q + a_r \le n \\ a_p + a_q + a_r - n, & \text{if } a_p + a_q + a_r > n \end{cases}$$

(Case(iii)) If $a_q + a_r > n$ then, $\mu_1 = a_q + a_r - n$ therefore,

$$\rho_1 = \begin{cases} a_p + a_q + a_r, & \text{if } a_p + a_q + a_r < 0 \\ a_p + a_q + a_r - n, & \text{if } 0 \le a_p + a_q + a_r \le 2n \\ a_p + a_q + a_r - 2n, & \text{if } a_p + a_q + a_r > 2n \end{cases}$$

Next, we let $a_p * a_q = \mu_2$ such that $(a_p * a_q) * a_r = \mu_2 * a_r = \rho_2$ therefore,

$$\rho_2 = \begin{cases} \mu_2 + a_r + n, & \text{if } \mu_2 + a_r < -n \\ \mu_2 + a_r, & \text{if } -n \le \mu_2 + a_r \le n \\ \mu + a_r - n, & \text{if } \mu_2 + a_r > n \end{cases} \quad (3)$$

Since $\mu_2 = a_p * a_q$, i.e.,

$$\mu_2 = \begin{cases} a_p + a_q + n, & \text{if } a_p + a_q < -n \\ a_p + a_q, & \text{if } -n \le a_p + a_q \le n \\ a_p + a_q - n, & \text{if } a_p + a_q > n \end{cases} \quad (4)$$

By substituting Eq. 4 in Eq. 3, we obtain the following cases which are defined below:

(Case(i)) If $a_p+a_q < -n$ then, $\mu_2 = a_p+a_q+n$ therefore,

$$\rho_2 = \begin{cases} a_p+a_q+a_r+2n, & \text{if } a_p+a_q+a_r < -2n \\ a_p+a_q+a_r+n, & \text{if } -2n \le a_p+a_q+a_r \le 0 \\ a_p+a_q+a_r, & \text{if } a_p+a_q+a_r > 0 \end{cases}$$

(Case(ii)) If $-n \le a_p+a_q \le n$ then, $\mu_2 = a_p+a_q$ therefore,

$$\rho_2 = \begin{cases} a_p+a_q+a_r+n, & \text{if } a_p+a_q+a_r < -n \\ a_p+a_q+a_r, & \text{if } -n \le a_p+a_q+a_r \le n \\ a_p+a_q+a_r-n, & \text{if } a_p+a_q+a_r > n \end{cases}$$

(Case(iii)) If $a_p+a_q > n$ then, $\mu_2 = a_p+a_q-n$ therefore,

$$\rho_2 = \begin{cases} a_p+a_q+a_r, & \text{if } a_p+a_q+a_r < 0 \\ a_p+a_q+a_r-n, & \text{if } 0 \le a_p+a_q+a_r \le 2n \\ a_p+a_q+a_r-2n, & \text{if } a_p+a_q+a_r > 2n \end{cases}$$

From all the above cases, we infer that $\rho_1 = \rho_2 \Rightarrow a_p * (a_q * a_r) = (a_p * a_q) * a_r$. Hence, X satisfies associative property. Next, 0 is an identity element such that for every $a_p \in [-n, n]$ we have $a_p * 0 = 0 * a_p = a_p$. Finally, for every $a_p \in X = [-n, n]$ their exist $(-a_p) \in X = [-n, n]$ such that $a_p*(-a_p) = (-a_p)*a_p = 0$. Hence, $X = [-n, n]$ is a group. □

Proposition 2 *Let $LT = \{t_{-n}, \ldots, t_0, t_1, \ldots, t_n\}$ be a finite SSLTS. Then LT is a group concerning binary operation '$*$' defined as:*

$$t_p * t_q = t_{p*q} = \begin{cases} t_{p+q+n}, & \text{if } p+q < -n \\ t_{p+q}, & \text{if } -n \le p+q \le n \\ t_{p+q-n}, & \text{if } p+q > n \end{cases}$$

for all $t_p, t_q \in LT$.

Proof We here give a detailed proof. Consider linguistic term set $LT = \{t_{-n}, \ldots, t_0, t_1, \ldots, t_n\}$ and let $t_p, t_q \in LT$ be any two arbitrary linguistic term. Define a binary operation '$*$' in the following aspect:

$$t_p * t_q = t_{p*q} = \begin{cases} t_{p+q+n}, & \text{if } p+q < -n \\ t_{p+q}, & \text{if } -n \le p+q \le n \\ t_{p+q-n}, & \text{if } p+q > n \end{cases}$$

Clearly, $t_p * t_q \in LT \ \forall \ t_p, t_q \in LT$. Hence, closure property holds. Next, let $t_p, t_q, t_r \in LT$ be a linguistic term. Let $t_q * t_r = t_{q*r} = \nu_1$ such that $t_p * (t_q * t_r) = t_p * \nu_1 = \tau_1$ therefore,

$$\begin{aligned} \tau_1 &= t_p * \nu_1 \\ &= t_p * t_{q*r} \\ &= \begin{cases} t_{p+(q*r)+n}, & \text{if } p+(q*r) < -n \\ t_{p+(q*r)}, & \text{if } -n \le p+(q*r) \le n \\ t_{p+(q*r)-n}, & \text{if } p+(q*r) > n \end{cases} \end{aligned} \tag{5}$$

since $\nu_1 = t_q * t_r = t_{q*r}$ i.e.,

$$\nu_1 = \begin{cases} t_{q+r+n}, & \text{if } q+r < -n \\ t_{q+r}, & \text{if } -n \le q+r \le n \\ t_{q+r-n}, & \text{if } q+r > n \end{cases} \tag{6}$$

By substituting Eq. 6 in Eq. 5, we obtain the following cases defined below :

(Case(i)) If $q+r < -n$ then, $\nu_1 = t_{q+r+n}$ therefore,

$$\begin{aligned} \tau_1 &= t_p * \nu_1 \\ &= t_p * t_{q*r} \\ &= \begin{cases} t_{p+q+r+2n}, & \text{if } p+q+r < -2n \\ t_{p+q+r+n}, & \text{if } -2n \le p+q+r \le 0 \\ t_{p+q+r}, & \text{if } p+q+r > 0 \end{cases} \end{aligned}$$

(Case(ii)) If $-n \le q+r \le n$ then, $\nu_1 = t_{q*r} = t_{q+r}$ therefore,

$$\begin{aligned} \tau_1 &= t_p * \nu_1 \\ &= t_p * t_{q*r} \\ &= \begin{cases} t_{p+q+r+n}, & \text{if } p+q+r < -n \\ t_{p+q+r}, & \text{if } -n \le p+q+r \le n \\ t_{p+q+r-n}, & \text{if } p+q+r > n \end{cases} \end{aligned}$$

(Case(iii)) If $q+r > n$ then, $\nu_1 = t_{q*r} = t_{q+r-n}$ therefore,

$$\begin{aligned} \tau_1 &= t_p * \nu_1 \\ &= t_p * t_{q*r} \\ &= \begin{cases} t_{p+q+r}, & \text{if } p+q+r < 0 \\ t_{p+q+r-n}, & \text{if } 0 \le p+q+r \le 2n \\ t_{p+q+r-2n}, & \text{if } p+q+r > 2n \end{cases} \end{aligned}$$

Next, let $t_p * t_q = t_{p*q} = \nu_2$ such that $(t_p * t_q) * t_r = \nu_2 * t_r = \tau_2$ therefore,

$$\begin{aligned} \tau_2 &= \nu_2 * t_r \\ &= t_{p*q} * t_r \end{aligned}$$

$$= \begin{cases} t_{(p*q)+r+n}, & \text{if } (p*q)+r < -n \\ t_{(p*q)+r}, & \text{if } -n \le (p*q)+r \le n \\ t_{(p*q)+r-n}, & \text{if } (p*q)+r > n \end{cases} \quad (7)$$

since $v_2 = t_p * t_q = t_{p*q}$ i.e.,

$$v_2 = \begin{cases} t_{p+q+n}, & \text{if } p+q < -n \\ t_{p+q}, & \text{if } -n \le p+q \le n \\ t_{p+q-n}, & \text{if } p+q > n \end{cases} \quad (8)$$

By substituting v_2 from Eq. 8 in Eq. 7, we obtain the following cases defined below:

(Case(i)) If $p+q < -n$ then, $v_2 = t_{p+q+n}$ therefore,

$$\begin{aligned} \tau_2 &= v_2 * t_r \\ &= t_{p+q+n} * t_r \\ &= \begin{cases} t_{p+q+r+2n}, & \text{if } p+q+r < -2n \\ t_{p+q+r+n}, & \text{if } -2n \le p+q+r \le 0 \\ t_{p+q+r}, & \text{if } p+q+r > 0 \end{cases} \end{aligned}$$

(Case(ii)) If $-n \le p+q \le n$ then, $v_2 = t_{p*q} = t_{p+q}$ therefore,

$$\begin{aligned} \tau_2 &= v_2 * t_r \\ &= t_{p+q} * t_r \\ &= \begin{cases} t_{p+q+r+n}, & \text{if } p+q+r < -n \\ t_{p+q+r}, & \text{if } -n \le p+q+r \le n \\ t_{p+q+r-n}, & \text{if } p+q+r > n \end{cases} \end{aligned}$$

(Case(iii)) If $p+q > n$ then, $v_2 = t_{p*q} = t_{p+q-n}$ therefore,

$$\begin{aligned} \tau_2 &= v_2 * t_r \\ &= t_{p+q-n} * t_r \\ &= \begin{cases} t_{p+q+r}, & \text{if } p+q+r < 0 \\ t_{p+q+r-n}, & \text{if } 0 \le p+q+r \le 2n \\ t_{p+q+r-2n}, & \text{if } p+q+r > 2n \end{cases} \end{aligned}$$

From all the above cases, we infer that $\tau_1 = \tau_2 \Rightarrow t_p * (t_q * t_r) = (t_p * t_q) * t_r$. Hence, LT satisfies associative property. Next, $t_0 \in LT$ is an identity element such that for every $t_p \in LT$ we have $t_p * t_0 = t_0 * t_p = t_p$. Finally, for every $t_p \in LT$ their exist $t_{-p} \in LT$ such that $t_p * t_{-p} = t_{-p} * t_p = 0$. Hence, LT is a group. □

Proposition 3 *Let $H = [\frac{-1}{2}, \frac{1}{2}]$ be a closed and bounded interval. Then H is a group concerning a binary operation '$\oplus$' given by.*

$$\alpha_p \oplus \alpha_q = \begin{cases} \alpha_p + \alpha_q, & \text{if } \alpha_p + \alpha_q \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \alpha_q - \text{round}(\alpha_p + \alpha_q), & \text{else} \end{cases}$$

Proof Let $H = [\frac{-1}{2}, \frac{1}{2}] = \left\{\alpha \mid \frac{-1}{2} \le \alpha \le \frac{1}{2}\right\}$ be a nonempty set. Let $\alpha_p, \alpha_q \in H$, then followed by the definition clearly, $\alpha_p \oplus \alpha_q \in H$. Next, let $\alpha_p, \alpha_q, \alpha_r \in H$. Consider $\alpha_q \oplus \alpha_r = \mu$ such that

$$\begin{aligned} \alpha_p \oplus (\alpha_q \oplus \alpha_r) &= \alpha_p \oplus \mu \\ &= \begin{cases} \alpha_p + \mu, & \text{if } \alpha_p + \mu \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \mu - \text{round}(\alpha_p + \mu), & \text{else} \end{cases} \end{aligned} \quad (9)$$

since,

$$\mu = \alpha_q \oplus \alpha_r = \begin{cases} \alpha_q + \alpha_r, & \text{if } \alpha_q + \alpha_r \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r), & \text{else} \end{cases}$$

By substituting value of μ in Eq. 9, we obtain the following cases define below:

(Case(i)) If $\alpha_q + \alpha_r \in (\frac{-1}{2}, \frac{1}{2})$ then, $\mu = \alpha_q + \alpha_r$ therefore,

$$\begin{aligned} \alpha_p \oplus (\alpha_q \oplus \alpha_r) &= \alpha_p \oplus \mu \\ &= \begin{cases} \alpha_p + \mu, & \text{if } \alpha_p + \mu \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \mu - \text{round}(\alpha_p + \mu), & \text{else} \end{cases} \\ &= \begin{cases} \alpha_p + \alpha_q + \alpha_r, & \text{if } \alpha_p + \alpha_q + \alpha_r \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \alpha_q + \alpha_r - \text{round}(\alpha_p + \alpha_q + \alpha_r), & \text{else} \end{cases} \end{aligned}$$

(Case(ii)) If $\mu \notin \left(\frac{-1}{2}, \frac{1}{2}\right)$ then, $\mu = \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r)$ therefore,

$$\alpha_p \oplus (\alpha_q \oplus \alpha_r) = \alpha_p \oplus \mu$$

$$= \begin{cases} \alpha_p + \mu, & \text{if } \alpha_p + \mu \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \mu - \text{round}(\alpha_p + \mu), & \text{else} \end{cases}$$

$$= \begin{cases} \alpha_p + \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r), & \text{if } \alpha_p + \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r) \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r) - \\ \quad \text{round}\left(\alpha_p + \alpha_q + \alpha_r - \text{round}(\alpha_q + \alpha_r)\right), & \text{else} \end{cases}$$

Similarly, we can have $(\alpha_p \oplus \alpha_q) \oplus \alpha_r$ and $\alpha_p \oplus (\alpha_q \oplus \alpha_r) = (\alpha_p \oplus \alpha_q) \oplus \alpha_r$ and thus associative property holds. Next, $0 \in [\frac{-1}{2}, \frac{1}{2}]$ is an identity element and finally for every $\alpha_p \in [\frac{-1}{2}, \frac{1}{2}]$ their exist $(-\alpha_p) \in [\frac{-1}{2}, \frac{1}{2}]$ such that $\alpha_p \oplus (-\alpha_p) = (-\alpha_p) \oplus \alpha_p = 0$. Hence, H is a group under a binary operation '$\oplus$'. □

It is known that the direct product of groups is itself forms a group (Gallian 2012). Therefore, in this direction, the underlying theorem gives an elementary proof of the 2-tuple linguistic term set to be a group.

Theorem 1 *Let $LT = \{t_{-n}, \ldots, t_0, t_1, \ldots, t_n\}$ be the SSLTS with odd cardinality $2n+1$ and let $\alpha_p \in [\frac{-1}{2}, \frac{1}{2}]$ be the numeric value representing the value of symbolic translation corresponding a linguistic term t_p. Then $\overline{LT} \equiv LT \times [\frac{-1}{2}, \frac{1}{2}]$ is a direct group under a binary operation '$\circ$' defined as:*

$$(t_p, \alpha_p) \circ (t_q, \alpha_q) = (t_p \circ t_q, \alpha_p \circ \alpha_q)$$

where,

$$t_p \circ t_q = t_p * t_q$$

$$= t_{p*q}$$

$$= \begin{cases} t_{p+q+n}, & \text{if } p+q < -n \\ t_{p+q}, & \text{if } -n \le p+q \le n \\ t_{p+q-n}, & \text{if } p+q > n \end{cases}$$

and

$$\alpha_p \circ \alpha_q = \alpha_p \oplus \alpha_q$$

$$= \begin{cases} \alpha_p + \alpha_q, & \text{if } \alpha_p + \alpha_q \in \left(\frac{-1}{2}, \frac{1}{2}\right) \\ \alpha_p + \alpha_q - \text{round}(\alpha_p + \alpha_q), & \text{else} \end{cases}$$

where $t_p, t_q \in LT$ and $\alpha_p, \alpha_q \in [\frac{-1}{2}, \frac{1}{2}]$.

Proof Let $\overline{LT} = \left\{(t_p, \alpha_p) \,|\, t_p \in LT \text{ and } \alpha_p \in [\frac{-1}{2}, \frac{1}{2}]\right\}$ be a nonempty set of 2-tuple linguistic terms. From the definition $(t_p, \alpha_p) \circ (t_q, \alpha_q) = (t_p \circ t_q, \alpha_p \circ \alpha_q) = (t_p * t_q, \alpha_p \oplus \alpha_q) \in \overline{LT} \,\forall\, (t_p, \alpha_p), (t_q, \alpha_q) \in \overline{LT}$, so closure property holds. Next associative property is followed by using Propositions 2 and 3 such that $(t_p, \alpha_p) \circ \left((t_q, \alpha_q) \circ (t_r, \alpha_r)\right) = \left((t_p, \alpha_p) \circ (t_q, \alpha_q)\right) \circ (t_r, \alpha_r)$. Next $(t_0, 0) \in \overline{LT}$ is an identity element and finally for every $(t_p, \alpha_p) \in \overline{LT}$ their exist $(t_{-i}, (-\alpha_p)) \in \overline{LT}$ such that

$$(t_p, \alpha_p) \circ (t_{-p}, (-\alpha_p)) = (t_{-p}, (-\alpha_p)) \circ (t_p, \alpha_p) = (t_0, 0)$$

Hence, $\overline{LT}$ is a group. □

Next, we present the main result of this section.

Theorem 2 *Let $LT = \{t_{-n}, \ldots, t_0, \ldots, t_{-n}\}$ be the finite SSLTS and let $x \in [-n, n]$ a numeric value supporting the outcome of an operation of symbolic aggregation. Then a 2-tuple linguistic direct group, $\overline{LT}$ and $[-n, n]$ are isomorphic.*

Proof Consider the mapping,

$\phi : [-n, n] \to \overline{LT}$ defined by
$\phi(x_p) = (t_p, \alpha_p)$

with,

$$\begin{cases} t_p, & p = \text{round}(x_p), \\ \alpha_p = x_p - p, & \alpha_p \in [\frac{-1}{2}, \frac{1}{2}]. \end{cases}$$

Let $x_p, x_q \in [-n, n]$ such that $\phi(x_p) = \phi(x_q) \Leftrightarrow (t_p, \alpha_p) = (t_q, \alpha_q) \Leftrightarrow (t_p, x_p - p) = (t_q, x_q - q) \Leftrightarrow t_p = t_q, x_p - p = x_q - q \Leftrightarrow p = q,\ x_p - p = x_q - q \Leftrightarrow x_p = x_q$. Therefore, ϕ is well-defined and one-one. ϕ is a onto function. This is easy to see that for every $(t_p, \alpha_p) \in \overline{LT}$ their exist $x_p \in [-n, n]$ such that,

$$\phi(x_p) = (t_p, \alpha_p) \text{ where } \begin{cases} t_p, & p = \text{round}(x_p), \\ \alpha_p = x_p - p, & \alpha_p \in [\frac{-1}{2}, \frac{1}{2}]. \end{cases}$$

Next, we proceed to show ϕ is homomorphic. Let $x_p,\ x_q \in [-n, n]$ such that,

$$\phi(x_p) = (t_p, \alpha_p), \text{ with } \begin{cases} t_p, & p = \text{round}(x_p), \\ \alpha_p = x_p - p, & \alpha_p \in [\frac{-1}{2}, \frac{1}{2}]. \end{cases}$$

$$\phi(x_q) = (t_q, \alpha_q), \text{ with } \begin{cases} t_q, & q = \text{round}(x_q), \\ \alpha_q = x_q - q, & \alpha_q \in [\frac{-1}{2}, \frac{1}{2}]. \end{cases}$$

$$\phi(x_p * x_q) = (t_r, \alpha_r), \text{ with } \begin{cases} t_r, & r = \text{round}(x_p * x_q) \\ \alpha_r = (x_p * x_q) - r, & \alpha_r \in [\frac{-1}{2}, \frac{1}{2}]. \end{cases}$$

We claim $\phi(x_p * x_q) = \phi(x_p) \circ \phi(x_q)$, i.e., to show $t_r = t_p * t_q$ and $\alpha_r = \alpha_p \oplus \alpha_q$, via Algorithm 1.

Algorithm 1 defines a decision rule to represent the homomorphic relation between the groups, considering all the possible cases such that

$$\begin{aligned} \phi(x_p * x_q) &= (t_r, \alpha_r) \\ &= (t_p * t_q, \alpha_p \oplus \alpha_q) \\ &= (t_p \circ t_q, \alpha_p \circ \alpha_q) \\ &= (t_p, \alpha_p) \circ (t_q, \alpha_q) \\ &= \phi(x_p) \circ \phi(x_q). \end{aligned}$$

Hence, ϕ is an isomorphism.

Algorithm 1 Decision rule to represent the homomorphic relation

(Case(i))

Require: $p + q < -n$, then, followed by the definition we have, $t_p * t_q = t_{p+q+n}$.

1: **if** $p + q + n = r$ **then**
2: $t_r = t_{p+q+n}$,
3: $\alpha_r = \alpha_p \oplus \alpha_q$.
4: **else**
5: t_r is represented as t_m, where $m \leftarrow r - 1$,
6: **if** $m \in (-n, n)$ **then**
7: $t_m = t_{p+q+n}$,
8: $\alpha_r = \alpha_p \oplus \alpha_q$.
9: **end if**
10: **if** $m \leq -n$ **then**
11: $m = r - 1 + n$ such that $t_m = t_{p+q+n} = t_r$,
12: $\alpha_r = \alpha_p \oplus \alpha_q$.
13: **end if**
14: **end if**

(Case(ii))

Require: $-n < p + q < n$, then, followed by the definition we have $t_p * t_q = t_{p+q}$.

1: **if** $r = p + q$ **then**
2: $t_r = t_{p+q}$,
3: $\alpha_r = \alpha_p \oplus \alpha_q$.
4: **else**
5: t_r is represented as t_m, where $m \leftarrow r + 1$ or $m \leftarrow r - 1$ such that
6: $t_r = t_m$,
7: $\alpha_r = \alpha_p \oplus \alpha_q$.
8: **end if**

(Case(ii))

Require: $p + q > n$, then, followed by the definition we have $t_p * t_q = t_{p+q-n}$.

1: **if** $p + q - n = r$ **then**
2: $t_r = t_{p+q-n}$,
3: $\alpha_r = \alpha_p \oplus \alpha_q$.
4: **else**
5: t_r is represented as t_m, where $m \leftarrow r + 1$,
6: **if** $m \in (-n, n)$ **then**
7: $t_m = t_{p+q-n}$,
8: $\alpha_r = \alpha_p \oplus \alpha_q$.
9: **end if**
10: **if** $m \geq n$ **then**
11: $m = r + 1 - n$ such that $t_m = t_{p+q-n} = t_r$,
12: $\alpha_r = \alpha_p \oplus \alpha_q$.
13: **end if**
14: **end if**

(Case(iv))

1: **if** $p + q = -n$ or $p + q = n$ **then**
2: $r \leftarrow n$ or $-n$ such that $t_r = t_{p+q}$ or $t_r = t_{-(p+q)}$,
3: $\alpha_r = \alpha_p \oplus \alpha_q$.
4: **end if**

□

It is requisite to view that the aforementioned algorithm guarantees that the mapping $x \rightarrow (t, \alpha)$ is an operation preserving map under a defined binary operation. Moreover, the time complexity of the proposed algorithm is linear, which has a practical advantage in laying the foundations for 2-tuple linguistic group homomorphism in several real-life problems.

In particular, if there exists an isomorphism of a group $\overline{LT}$ to itself, i.e., $\overline{LT} \rightarrow \overline{LT}$ then it is termed as an automorphism of a 2-tuple linguistic group.

We now consider the following illustration to understand the working of the algorithm as mentioned above and further combining the result to indicate that the mapping $x \rightarrow (t, \alpha)$ is an isomorphism defined from $[-n, n]$ to $\overline{LT}$ under binary operation previously stated.

Example 1 Consider a mapping $\phi : [-3, 3] \to \overline{LT}$ such that $\phi(x_p) = (t_p, \alpha_p)$ where $t_p \in \{t_{-3}, t_{-2}, t_{-1}, t_0, t_1, t_2, t_3\}$ be the SSLTS and $\alpha_p \in \left[\frac{-1}{2}, \frac{1}{2}\right]$. Clearly, the defined mapping is bijective. Next, in order to establish isomorphism between the groups we need to preserve the group operation. In this direction, we shall claim $\phi(x_p * x_q) = \phi(x_p) \circ \phi(x_q)$, i.e., to show $t_r = t_p * t_q$ and $\alpha_r = \alpha_p \oplus \alpha_q$.

For this, we use Algorithm 1. Since the order of both the groups are infinite, therefore, we have constructed a MATLAB code based on the aforementioned algorithm to generate numbers randomly to perform further routine computations. We present in Table 1, the list of cases to accomplish the result.

From Table 1, we can infer that the mapping is group operation preserving, i.e., $\phi(x_p * x_q) = \phi(x_p) \circ \phi(x_q)$ and therefore, $[-3, 3]$ is isomorphic to $\overline{LT}$.

Remark 3 It can be noted here that in order to simplify the calculations and understand the result described above, we have taken $n = 3$. However, the procedure is analogous for any integral value of n.

4 Application of 2-tuple linguistic group in bipolar graphs

In the literature, (Zhang 1994, 1989) presented the notion of bipolar fuzzy sets. It is visualized that the extent of membership degree for a bipolar fuzzy set lies within the range $[-1, 1]$, whereby it is considered as an extrapolation of fuzzy sets. Here, the element with 0 membership degree is considered insignificant to the associated property. In contrast, the elements whose membership degree lying in the range $(0, 1]$ represent the element fulfills the property, unlike the elements whose degree of membership lying within the range $[-1, 0)$ represent the element is somewhat fulfilling the implicit counter-property. Nevertheless, bipolar information plays a significant role in a wide-ranging multiagent decision analysis where both-sided judgemental thinking is involved, and therefore, this domain has pervasively inspired several researchers across the globe. Recently, (Akram 2011; Akram and Dudek 2012; Akram and Karunambigai 2011) has put forward the impression of bipolar fuzzy graphs to amplify the purview of bipolar fuzzy sets within the construction of a structure of fuzzy graphs.

The amalgamation of linguistic information with the existing notion of fuzzy sets has allegedly played a pivotal role in handling the uncertainty persisting within the problem. Therefore, inspired by the work of Akram (2011) in the present section, we bring forth the foundation of 2-tuple linguistic bipolar set (2TLBS) and, after that, discuses the basic algebra on the set. We later make headway to give the application of 2-tuple linguistic groups by developing the notion of 2-tuple linguistic bipolar graphs (2TLBG) and further establishing the graph isomorphic relation.

4.1 The 2-tuple linguistic bipolar set

It is well known that the popularly known 2-tuple linguistic representation is ideally introduced by Herrera and Martínez for the linguistic term set $S = \{t_p \mid p = 0, 1, \ldots, n\}$ primarily based upon the symbolic translation. Since the traditional 2-tuple model is capable of effectively handling linguistic information without any loss. As a consequence, we will first introduce the notion of 2TLBS based on the 2-tuple model using predefined SSLTS $LT = \{t_p \mid p = -n, \ldots, 0, \ldots, n\}$ followed by some basic operations on the set which are required to establish the conception of 2TLBG given in the subsequent part of the section.

Definition 12 Let $X = [-n, n]$ be any arbitrary nonempty set and let $LT = \{t_p \mid p = -n, \ldots, 0, \ldots, n\}$ be a SSLTS. Then, we define 2-tuple linguistic bipolar set on X as follows:

$$\overline{\mathbf{C}} = \{(x, \phi^+(|x|), \phi^-(-|x|)) \mid x \in X\}$$

with $\phi^+ : X \to LT^+ \times \left[\frac{-1}{2}, \frac{1}{2}\right)$ such that $\phi^+(x) = (t_{p^+}, \alpha)$ where $t_{p^+} \in LT^+ = \{t_p \mid p = 0, 1, \ldots, n\}$ and

Table 1 Computations for randomly generated numbers by using Algorithm 1

x_p	x_q	$\phi(x_p * x_q)$	$\phi(x_p) \circ \phi(x_q)$
1.8	2.8	$(t_2, -0.4)$	$(t_2, -0.2) \circ (t_3, -0.2) = (t_2, -0.4)$
0.8	−2.4	$(t_{-1}, 0.4)$	$(t_1, -0.2) \circ (t_{-2}, -0.4) = (t_{-1}, 0.4)$
−2.8	−2.4	$(t_{-2}, -0.2)$	$(t_{-3}, 0.2) \circ (t_{-2}, -0.4) = (t_{-2}, -0.2)$
−0.4	2.5	$(t_3, 0.1)$	$(t_0, -0.4) \circ (t_3, -0.5) = (t_3, 0.1)$
2.0	−1.5	$(t_0, 0.5)$	$(t_2, 0) \circ (t_{-2}, 0.5) = (t_0, 0.5)$
1.5	1.5	$(t_1, 0)$	$(t_2, -0.5) \circ (t_2, -0.5) = (t_1, 0)$
−2.4	1.9	$(t_0, -0.5)$	$(t_{-2}, -0.4) \circ (t_2, -0.1) = (t_0, -0.5)$

$\phi^- : X \to LT^- \times \left(\frac{-1}{2}, \frac{1}{2}\right]$ such that $\phi^-(x) = (t_{p^-}, \alpha)$ where $t_{p^-} \in LT^- = \{t_p \mid p = -n, \ldots, -1, 0\}$.

Here, t_{p^+} represents one of the positive linguistic terms from the set LT^+ whereas t_{p^-} represents the counter linguistic term belonging to the LT^- and α indicates a numerical value that captures the uncertainty about a corresponding linguistic term. Consequently, both $\phi^+(x)$ and $\phi^-(x)$ represent the positive and negative 2-tuple representation based on the linguistic term, respectively.

For notational simplicity, we use $\overline{\mathbf{C}} = (\phi^+, \phi^-)$ to refer 2TLBS.

Example 2 Let $X = [-2, 2]$ and let

$$\begin{aligned} LT = \{&t_{-2} = \text{“Very poor (VP)”}, t_{-1} = \text{“Poor (P)”}, \\ &t_0 = \text{“Fair (F)”}, t_1 = \text{“Good (G)”}, t_2 \\ &= \text{“Very Good (VG)”}\} \end{aligned}$$

be predefined SSLTS. Then,

$$\begin{aligned} \overline{\mathbf{C}} &= \{(-1.2, \phi^+(1.2), \phi^-(-1.2)), (-0.5, \phi^+(0.5), \\ &\quad \phi^-(-0.5)), (1.8, \phi^+(1.8), \phi^-(-1.8))\} \\ &= \{(-1.2, (G, 0.2), (P, -0.2)), (-0.5, (G, -0.5), \\ &\quad (P, 0.5)), (1.8, (VG, -0.2), (VP, 0.2))\} \end{aligned}$$

is the 2TLBS for given set X.

Definition 13 Suppose $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}}, \overline{\mathbf{D}} = (\phi^+_{\mathbf{D}}, \phi^-_{\mathbf{D}})$ are any two 2TLBS defined on set X. Then the following operations for set $\overline{\mathbf{C}}$ and $\overline{\mathbf{D}}$ hold:

1. $\left(\overline{\mathbf{C}} \cup \overline{\mathbf{D}}\right)(x) = \left(\phi^+_{\mathbf{C}\cup\mathbf{D}}, \phi^-_{\mathbf{C}\cup\mathbf{D}}\right)$ where $\phi^+_{\mathbf{C}\cup\mathbf{D}} = \max(\phi^+_{\mathbf{C}}, \phi^+_{\mathbf{D}})$ and $\phi^-_{\mathbf{C}\cup\mathbf{D}} = \min(\phi^-_{\mathbf{C}}, \phi^-_{\mathbf{D}})$.
2. $\left(\overline{\mathbf{C}} \cap \overline{\mathbf{D}}\right)(x) = \left(\phi^+_{\mathbf{C}\cap\mathbf{D}}, \phi^-_{\mathbf{C}\cap\mathbf{D}}\right)$ where $\phi^+_{\mathbf{C}\cap\mathbf{D}} = \min(\phi^+_{\mathbf{C}}, \phi^+_{\mathbf{D}})$ and $\phi^-_{\mathbf{C}\cap\mathbf{D}} = \max(\phi^-_{\mathbf{C}}, \phi^-_{\mathbf{D}})$.

It is worth noting that the necessary operational laws defined in the above definition are valid as the comparison of two 2-tuple linguistic variables is plausible. However, the other set operations can be easily deduced for 2TLBSs by combining the operations existing for the bipolar fuzzy sets and the operations of 2-tuple linguistic variables.

Next, based on the result proved in the former section that the 2-tuple linguistic set $\overline{\mathrm{LT}} \equiv \mathrm{LT} \times [\frac{-1}{2}, \frac{1}{2}]$ is a direct group under a binary operation '$\circ$' we proceed further to review the following definition required for this paper.

Definition 14 Assume X to be a nonempty set and $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ be any 2TLBS. Then, a mapping $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}}) : X \times X \to \overline{LT} \times \overline{LT}$ is known as 2-tuple linguistic bipolar relation on set X such that $\phi^+_{\mathbf{C}} \in LT^+ \times [\frac{-1}{2}, \frac{1}{2})$ and $\phi^-_{\mathbf{C}} \in LT^- \times (\frac{-1}{2}, \frac{1}{2}]$.

Definition 15 Let X be any arbitrary set and let $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ and $\overline{\mathbf{D}} = (\phi^+_{\mathbf{D}}, \phi^-_{\mathbf{D}})$ be any two 2TLBSs on set X. If $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ is a 2-tuple linguistic bipolar relation. Then, $\overline{\mathbf{C}}$ can be considered as a 2-tuple linguistic bipolar relation on $\overline{\mathbf{D}}$ satisfying the subsequent condition:

$$\phi^+_{\mathbf{C}}(c, d) \leq \min(\phi^+_{\mathbf{D}}(c), \phi^+_{\mathbf{D}}(d))$$

and

$$\phi^-_{\mathbf{C}}(c, d) \geq \max(\phi^-_{\mathbf{D}}(c), \phi^-_{\mathbf{D}}(d)) \quad \text{for all} \quad c, d \in X.$$

Note 4 We call 2-tuple linguistic bipolar relation $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ on set X as symmetric if $\phi^+_{\mathbf{C}}(c, d) = \phi^+_{\mathbf{C}}(d, c)$ and $\phi^-_{\mathbf{C}}(c, d) = \phi^-_{\mathbf{C}}(d, c)$ for all $c, d \in X$.

4.2 The 2-tuple linguistic bipolar graphs

With the above concept of 2TLBS, we herein propose 2TLBGs and further consider an application of group operations and isomorphic relation already proved for the linguistic term set of 2-tuples to 2TLBGs by developing an isomorphic relation between the graphs.

In the classical graph theory (Harary 1972), a graph is considered as an ordered pair $T = (V, E)$ where V and E represent a set of vertices and edges of T, respectively. Analogously, we define the 2-tuple linguistic bipolar graphs based on SSLTS.

It is noteworthy to take $\mathbf{T}^*$ as a notation for 2TLBG, and T is a notation for crisp graphs throughout this paper.

Definition 16 Let $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ and $\overline{\mathbf{D}} = (\phi^+_{\mathbf{D}}, \phi^-_{\mathbf{D}})$ be any two 2TLBSs on set V and $E \subseteq V \times V$, respectively. Then, we define 2TLBG as $\mathbf{T}^* = (\overline{\mathbf{C}}, \overline{\mathbf{D}})$ such that $\phi^+_{\mathbf{D}}(\{c, d\}) \leq \min(\phi^+_{\mathbf{C}}(c), \phi^+_{\mathbf{C}}(d))$ and $\phi^-_{\mathbf{D}}(\{c, d\}) \geq \max(\phi^-_{\mathbf{C}}(c), \phi^-_{\mathbf{C}}(d))$ for all $\{c, d\} \in E$.

Here, we represent $\overline{\mathbf{C}}$ as 2-tuple linguistic bipolar vertex set of V, $\overline{\mathbf{D}}$ as 2-tuple linguistic bipolar edge set E, respectively. Also, we use cd as the notation for an element of edge set E, i.e., $\{c, d\} \equiv cd$.

Example 3 Consider a crisp graph $T = (V, E)$ such that $V = \{u, v, w\}$ and $E = \{uv, vw, wu\}$. Suppose $X = [-3, 3]$ and $LT = \{t_p \mid i = -3, -2, -1, 0, 1, 2, 3\}$ be predefined SSLTS. Let $\overline{\mathbf{C}} = (\phi^+_{\mathbf{C}}, \phi^-_{\mathbf{C}})$ and $\overline{\mathbf{D}} = (\phi^+_{\mathbf{D}}, \phi^-_{\mathbf{D}})$ be any two 2TLBSs defined by

	u	v	w
$\phi^{+}_{\overline{\mathbf{C}}}$	$(t_2,\ 0.4)$	$(t_2,\ -0.5)$	$(t_0,\ 0.3)$
$\phi^{-}_{\overline{\mathbf{C}}}$	$(t_{-2},\ -0.4)$	$(t_{-2},\ 0.5)$	$(t_0,\ -0.3)$
	uv	vw	wu
$\phi^{+}_{\overline{\mathbf{D}}}$	$(t_1,\ 0.4)$	$(t_0,\ 0.1)$	$(t_0,\ 0.2)$
$\phi^{-}_{\overline{\mathbf{D}}}$	$(t_{-1},\ -0.4)$	$(t_0,\ -0.1)$	$(t_0,\ -0.2)$

From routine calculations, we infer that $\mathbf{T}^* = (\overline{\mathbf{C}},\ \overline{\mathbf{D}})$ is a 2TLBG of T (as given in Fig. 1).

In the literature of classical graph theory (Harary 1972), graphs T_1 and T_2 are said to be isomorphic if there is a one-one and onto mapping between the sets of vertices of T_1 and T_2 such that any two arbitrary vertices v_1 and v_2 are adjoined in T_1 if and only if corresponding vertex $f(v_1)$ and $f(v_2)$ in graph T_2 are adjoined. If such an order-preserving bijective map exists between the two graphs, then we call the corresponding graph to be isomorphic, i.e., $T_1 \cong T_2$. On the other hand, graph automorphism is defined as an isomorphism of the graph itself, i.e., a well-defined mapping defined from the vertices of the given graph T to the vertices of T so that the following graph T is isomorphic with itself.

Based on the notion of isomorphism of crisp graphs, we shall now present the following definition to define an isomorphism of 2TLBGs considered to be a direct application of a 2-tuple linguistic group isomorphism.

Definition 17 Let $\mathbf{T_1}^* = (\overline{\mathbf{C_1}},\ \overline{\mathbf{D_1}})$ and $\mathbf{T_2}^* = (\overline{\mathbf{C_2}},\ \overline{\mathbf{D_2}})$ be any two 2TLBGs. A homomorphism $f : \mathbf{T_1}^* \to \mathbf{T_2}^*$ is a mapping $f : V_1 \to V_2$ satisfying the subsequent conditions:

1. $\phi^{+}_{\overline{\mathbf{C_1}}}(c_1) \leq \phi^{+}_{\overline{\mathbf{C_2}}}(f(c_1)), \phi^{-}_{\overline{\mathbf{C_1}}}(c_1) \geq \phi^{-}_{\overline{\mathbf{C_2}}}(f(c_1))$;
2. $\phi^{+}_{\overline{\mathbf{D_1}}}(c_1d_1) \leq \phi^{+}_{\overline{\mathbf{D_2}}}(f(c_1)f(d_1))$ and $\phi^{-}_{\overline{\mathbf{D_1}}}(c_1d_1) \geq \phi^{-}_{\overline{\mathbf{D_2}}}(f(c_1)f(d_1))$.

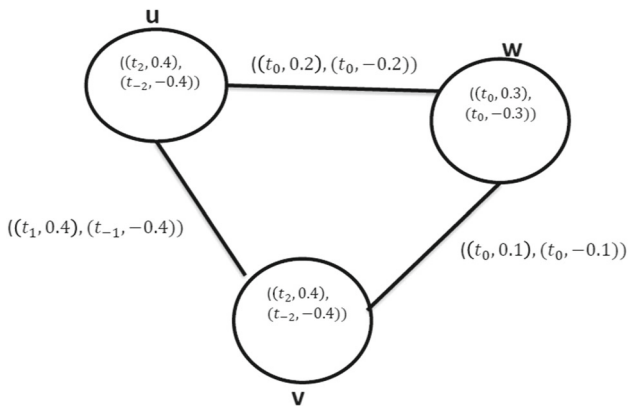
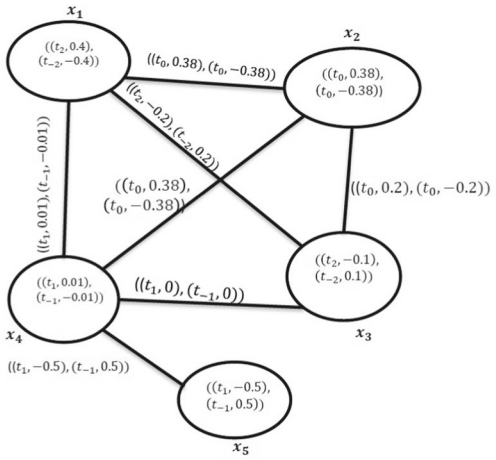
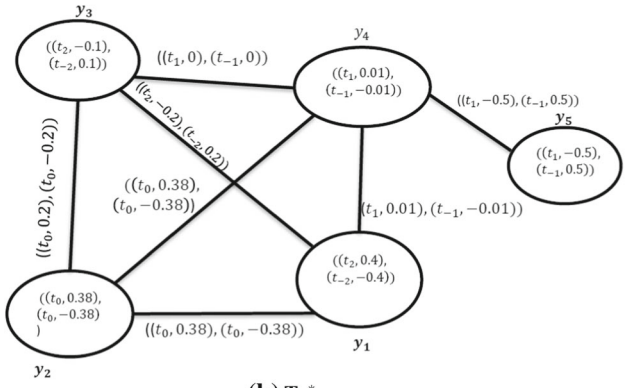


Fig. 1 A 2-tuple linguistic bipolar graph $\mathbf{G}^*$



(a) $\mathbf{T_1}^*$



(b) $\mathbf{T_2}^*$

Fig. 2 A 2-tuple linguistic bipolar graphs $\mathbf{T_1}^*$, $\mathbf{T_2}^*$

for all $c_1 \in V_1,\ c_1d_1 \in E_1$.

Definition 18 Let $\mathbf{T_1}^* = (\overline{\mathbf{C_1}},\ \overline{\mathbf{D_1}})$ and $\mathbf{T_2}^* = (\overline{\mathbf{C_2}},\ \overline{\mathbf{D_2}})$ be any two 2TLBGs. An isomorphism $f : \mathbf{T_1}^* \to \mathbf{T_2}^*$ is a bijective map $f : V_1 \to V_2$ satisfying the following conditions:

1. $\phi^{+}_{\overline{\mathbf{C_1}}}(c_1) = \phi^{+}_{\overline{\mathbf{C_2}}}(f(c_1)), \phi^{-}_{\overline{\mathbf{C_1}}}(c_1) = \phi^{-}_{\overline{\mathbf{C_2}}}(f(c_1))$;
2. $\phi^{+}_{\overline{\mathbf{D_1}}}(c_1d_1) = \phi^{+}_{\overline{\mathbf{D_2}}}(f(c_1)f(d_1))$ and $\phi^{-}_{\overline{\mathbf{D_1}}}(c_1d_1) = \phi^{-}_{\overline{\mathbf{D_2}}}(f(c_1)f(d_1))$.

for all $c_1 \in V_1,\ c_1d_1 \in E_1$.

We can observe that the above mappings are well-defined since the 2-tuple linguistic set forms a group under a given binary operation, and also, ϕ is an isomorphism. Consequently, the graphical structures of an isomorphic groups are similar.

Example 4 Suppose $X = [-3,\ 3]$ and $LT = \{t_p \mid p = -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3\}$ be predefined SSLTS. Consider 2TLBGs $\mathbf{T_1}^*$ and $\mathbf{T_2}^*$ (as shown in Fig. 2).

Therefore, by making routine computations we can observe that the two graphs are isomorphic, i.e., $\mathbf{T_1}^* \cong \mathbf{T_2}^*$.

5 Conclusion

In this work, we have developed a formal method that shows the intervals $[-n, n]$ and $[\frac{-1}{2}, \frac{1}{2}]$, and LT forms a group with respect to a binary operation in a crisp sense. Then, we have proved that the 2-tuple linguistic term set i.e., $\overline{LT} \equiv LT \times [\frac{-1}{2}, \frac{1}{2}]$ is a direct group under the defined operation '$\circ$' and later on developed an isomorphic relation between the 2-tuple linguistic direct group and the interval $[-n, n]$. An application in terms of bipolar linguistic graphs is also given to exhibit the effectiveness and practicality of the developed approach. In this connection, we initially defined the concept of 2TLBS, followed by the introduction of 2TLBGs. We have also established the isomorphic relation between the 2TLBGs, which directly follows from the 2-tuple linguistic group isomorphism. This is because when the two groups are isomorphic, they are abstractly similar and have the same graphical structures.

Nevertheless, the findings predicted in the paper are described using the SSLTS, the terms of which are uniformly and symmetrically positioned on either side of the middle term. We tend to envision that the research presented in the direction of the 2-tuple linguistic term set is to commence the theory of 2-tuple linguistic groups by utilizing classical group theory-based definitions and results. We believe that the findings proposed in this study will give rise to a new algebraic view of the 2-tuple model and can be seen as a significant step toward enhancing the importance of classical group theory in the uncertainty domain.

Compliance with Ethical Standards

Conflict of Interest Authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Akgül M (1988) Some properties of fuzzy groups. J Math Anal Appl 133(1):93–100

Akram M (2011) Bipolar fuzzy graphs. Inf Sci 181(24):5548–5564

Akram M, Dudek WA (2012) Regular bipolar fuzzy graphs. Neural Comput Appl 21:197–205. https://doi.org/10.1007/s00521-011-0772-6

Akram M, Karunambigai MG (2011) Metric in bipolar fuzzy graphs. World Appl Sci J 14(12):1920–1927

Anthony JM, Sherwood H (1979) Fuzzy groups redefined. J Math Anal Appl 69(1):124–130

Asaad M, Abou-Zaid S (1996) A contribution to the theory of fuzzy subgroups. Fuzzy Sets Syst 77(3):355–369

Bhattacharya P (1987) Fuzzy subgroups: some characterizations. J Math Anal Appl 128(1):241–252

Bhattacharya P, Mukherjee NP (1987) Fuzzy groups: some group theoretic analogs II. Inf Sci 41(1):77–91

Bonissone PP, Decker KS (1986) Selecting uncertainty calculi and granularity: an experiment in trading-off precision and complexity. Mach Intell Pattern Recognit 4:217–247

Bordogna G, Fedrizzi M, Pasi G (1997) A linguistic modeling of consensus in group decision making based on OWA operators. IEEE Trans Syst Man Cybern-Part A: Syst Humans 27(1):126–133

Das PS (1981) Fuzzy groups and level subgroups. J Math Anal Appl 84(1):264–269

De Luca A, Termini S (1972) Algebraic properties of fuzzy sets. J Math Anal Appl 40(2):373–386

Degang C, Jiashang J, Congxin W, Tsang EC (2005) Some notes on equivalent fuzzy sets and fuzzy subgroups. Fuzzy Sets Syst 152(2):403–409

Degani R, Bortolan G (1988) The problem of linguistic approximation in clinical decision making. Int J Approx Reason 2(2):143–162

Delgado M, Verdegay JL, Vila MA (1992) Linguistic decision making models. Int J Intell Syst 7(5):479–492

Delgado M, Herrera F, Herrera-Viedma E, Martínez L (1998) Combining numerical and linguistic information in group decision making. Inf Sci 107(1–4):177–194

Demirci M, Recasens J (2004) Fuzzy groups, fuzzy functions and fuzzy equivalence relations. Fuzzy Sets Syst 144(3):441–458

Dietrich H, Wilson JB (2019) Isomorphism testing of groups of cube-free order. J Algebr 545:174–197. https://doi.org/10.1016/j.jalgebra.2019.02.008

Dubois D, Prade H (1980) Fuzzy Sets and Systems: Theory and Applications. Kluwer Academic, New York

Gallian J (2012) Contemporary abstract algebra. Nelson Education

Gupta KC, Sarma BK (1999) Nilpotent fuzzy groups. Fuzzy Sets Syst 101(1):167–176

Harary F (1972) Graph Theory, 3rd edn. Addison-Wesley, Reading

Herrera F, Martínez L (2000) A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans Fuzzy Syst 8(6):746–752

Herrera F, Martínez L (2001) The 2-tuple linguistic computational model advantages of its linguistic description, accuracy and consistency. Int J Uncertain Fuzziness Knowl-Based Syst 9(1):33–48

Herrera F, Herrera-Viedma E, Verdegay JL (1995) A sequential selection process in group decision making with linguistic assessment. Inf Sci 85(4):223–239

Herrera F, Herrera-Viedma E, Verdegay JL (1996) Direct approach processes in group decision making using linguistic OWA operators. Fuzzy Sets Syst 79(2):175–190

Kahrobaei D, Anshel M (2010) Applications of group theory in cryptography. Int J Pure Appl Math 58:21–23

Kim JG (1997) Some characterizations of fuzzy subgroups. Fuzzy Sets Syst 87(2):243–249

Kleiner I (1986) The evolution of group theory: A brief survey. Math Mag 59(4):195–215

Klir GJ, Yuan B (1995) Fuzzy Sets an Fuzzy Logic: Theory and Applications, vol 4. Prentice-Hall PTR, USA

Makamba BB (1992) Direct products and isomorphism of fuzzy subgroups. Inf Sci 65(1–2):33–43

Malhotra T, Gupta A, Singh A (2019) Methodology for interval-valued matrix games with 2-tuple fuzzy linguistic information. In: International Conference on Numerical Computations: Theory and Algorithms (Springer), pp. 154–168

Mukherjee NP, Bhattacharya P (1984) Fuzzy normal subgroups and fuzzy cosets. Inf Sci 34(3):225–239

Mukherjee NP, Bhattacharya P (1986) Fuzzy groups: Some group-theoretic analogs. Inf Sci 39(3):247–267
Myasnikov A, Shpilrain V, Ushakov A (2008) Group-Based Cryptography. Birkhauser, Switzerland
Ovchinnikov S (1995) Continuous fuzzy groups. Fuzzy Sets Syst 76(2):253–257
Ray S (1992) Isomorphic fuzzy groups. Fuzzy Sets Syst 50(2):201–207
Ray AK (1999) On product of fuzzy subgroups. Fuzzy Sets Syst 105(1):181–183
Rodríguez-Lara BM, El-Ganainy R, Guerrero J (2018) Symmetry in optics and photonics: a group theory approach. Sci Bull 63(4):244–251
Rosenfeld A (1971) Fuzzy groups. J Math Anal Appl 35(3):512–517
Schweizer B, Sklar A (1960) Statistical metric spaces. Pacific J. Math. 10(1):313–334
Sherwood H (1983) Product of fuzzy subgroups. Fuzzy Sets Syst 11(1–3):79–89
Tărnăuceanu M, Bentea L (2008) On the number of fuzzy subgroups of finite abelian groups. Fuzzy Sets Syst 159(9):1084–1096
Tong RM, Bonissone PP (1980) A linguistic approach to decision making with fuzzy sets. IEEE Trans Syst Man Cybern 10(11):716–723
Tu L, Wenxiang G (1994) Abelian fuzzy group and its properties. Fuzzy Sets Syst 64(3):415–420
Xu Z (2005) Deviation measures of linguistic preference relations in group decision making. Omega 33(3):249–254
Yager RR (1981) A new methodology for ordinal multiobjective decisions based on fuzzy sets. Decis Sci 12(4):589–600. https://doi.org/10.1111/j.1540-5915.1981.tb00111.x
Yuan X, Lee ES (2004) Fuzzy group based on fuzzy binary operation. Comput Math Appl 47(4–5):631–641
Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-I. Inf Sci 8:199–249
Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-II. Inf Sci 8(4):301–357
Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning-III. Inf Sci 9(1):43–80
Zadeh LA (1976) The linguistic approach and its application to decision analysis. In: Ho YC, Mitter SK (eds) Directions in Large-Scale Systems. Springer, Boston, pp 260–282
Zhang WR (1994) Bipolar fuzzy sets and relations: a computational framework for cognitive modeling and multiagent decision analysis. In: Proceedings of IEEE Conference pp. 305–309 https://doi.org/10.1109/IJCF.1994.375115
Zhang, WR (1998) Bipolar fuzzy sets. Proceedings of FUZZ-IEEE 835–840

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

JOURNAL PRESS INDIA
www.journalpressindia.com

VISION: Journal of Indian Taxation
Vol. 7 (1), Jan-Jun 2020, pp. 1-21
DOI: 10.17492/vision.v7i1.195702
https://www.journalpressindia.com
© 2020 Journal Press India

Growth Story NDA versus UPA: Is Special Status of Indian States Justified?

*K. V. Bhanu Murthy**

ABSTRACT

As a sequel to an earlier paper, this paper considers the growth patterns of SDP (State Domestic Product) amongst Special and non-special states. Using semi-log growth equations and a Convergence Index, the paper demonstrates that the state GDP of non-special states has grown at 13.86% per annum while GDP of Special States amounts to 14.20% (ACGR). The special states have converged with non-special states at the rate of 0.29% per annum. However, the pattern of convergence was prominent during UPA1 and UPA2 while during NDA1 and NDA2, Special states have receded from non-special states. Special states grew faster during NDA 2 regime (2014 onwards). Thus the conclusion is that special status is justified, especially during the recent NDA regime. The growth rate of former J&K state has been reasonably high, irrespective of whether it was NDA or UPA regime. This counters the argument of instability in J&K. During NDA2 former J&K States' growth was over 20% per annum for the period 2014-2016!!

***Keywords**: Special status; State domestic product; Growth patterns; J&K.*

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

1.0 Introduction

The underlying philosophy of special status runs from developmental and regional inequalities and idiosyncrasies to fiscal federalism. Hence, such a step cannot be justified only on political grounds. The justification for Special status also flows from the principle of 'affirmative action'. Although, 'affirmative action' has hitherto, been used, only in the context of caste-based reservations in jobs and education, the basic principle is such action by the State which attempts to reverse a past historic disadvantage. The grant of Special status, in this context could, therefore, be seen as 'affirmative action'.

**Professor, Department of USME, Delhi Technological University, Delhi, India (E-mail: bhanumurthykv@yahoo.com)*

If we assume that economic growth patterns of Indian states are a product of three factors:

- regional factors (historic disadvantages);
- political regimes like NDA vs. UPA governments; and
- the native growth propensity;

then it is important to combine these three factors and analyse the growth patterns of SGDP (State Gross Domestic Product) amongst special and non-special states, during the different political regimes of NDA and UPA governments, in the past financial years 1993-94 to 2015-16.

In this paper, as a sequel to the 'Trends in Own Revenue of Indian States' (Murthy, 2018a), we study the 'Growth Story NDA vs. UPA regimes' in terms of the trends in state domestic product of Indian states. The study is especially designed to enable a comparative study of the pattern of SDP growth of special category states vs. non-special category states. In this study, we have taken 20 non-special states and 11 special states. There are certain 'Criteria for Special Category Status'. The National Development Council (NDC), a body of the former Planning Commission, had granted special status. This status is granted to regions that have been in a disadvantaged position for long, in comparison with the rest of the country. In the case of J&K, it was a political decision. The implications of special status in J&K were that it gives them special rights and privileges regarding employment with the state government, acquisition of property in the state, settling in the state, and the right to scholarships and other forms of aid that the state government provides. Another angle to special status granted to 11 Indian States arises out of fiscal provisions. The motivation arises out of the constitutional imperative of fiscal support from the Centre to States. The underlying philosophy of special status runs from fiscal federalism to developmental and regional inequalities and idiosyncrasies, both strategic and locational.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

2.0 Literature Review

A general review of the performance of the Indian States would be in place. Firstly, what is state domestic product or SDP? What is the ranking of GSDP amongst the Indian States for the year 2019?

A general definition of GSDP is: "Gross State Domestic Product (GSDP) is defined as a measure, in monetary terms, of the volume of all goods and services produced within the boundaries of the state during a given period of time, accounted without duplication" (Central Statistical Office (CSO), National Accounts Statistics).

How is it calculated? The formula to calculate the components of GSDP is Y = C + I + G + X + Z. That stands for GNP = Consumption + Investment + Government + X (net exports, or imports minus exports) + Z (net income earned by domestic residents from overseas investments - net income earned by foreign residents from domestic investments).

Maharashtra has highest GSDP among 33 Indian States and Union Territories. As of year, 2016-17, Maharashtra contributes 14.69% of total India's GDP at current prices followed by Tamil Nadu (8.27%) (Table 1). India's most populous state Uttar Pradesh is at 3rd position with share of 8.13% (Central Statistical Office (CSO), National Accounts Statistics).

Table 1: Ranking of GSDP in 2019

Rank	State/Union territory	Data year
1	Maharashtra	2019–20 est.
2	Tamil Nadu	2019–20 est.
3	Karnataka	2019–20 est.
4	Uttar Pradesh	2019–20 est.

Source: Central Statistical Office.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Murthy (2015) explains that "The Indian tax system has successfully mobilized resources to finance administrative, welfare and developmental activities of public authorities. Besides being the main source of revenue for both Central and State Governments, it is an effective instrument to realise various socio-economic objectives of national policies".

Murthy (2018) establishes the *case for* special status by estimating that "they are converging, at a rate of 2.85% per annum, with non-special states. All growth rates are highly significant which leaves no room for ambiguity about the broad conclusion that Special Status has proved to be justified.*"*

Bhaskar and Gupta (2007) summarize the recent 'growth pattern of India (and) is set in the context of the parallel experience of China, the experience of poverty reduction…' (p.135), illuminating India's economic development.

Kohli (2006) has argued against the notion that growth picked up after pro-reform structural changes since 1991. He argues that there three reasons to believe that this was not the case. Firstly, he states that pro-business growth strategies started in 1980. Secondly, after 1991 actually, industrial growth suffered. Thirdly, the subsequent growth had adverse distributional political consequences. One of the consequences was uneven growth across regions.

Shand and Bhide (2000) examine variations in size, income and structural characteristics of Indian states. It analyses trends in Net State Domestic Product amongst Indian States from 1970-71 to 1995-96. However, the emphasis is on sectoral patterns in growth. It speaks of the role of industry, agriculture and service sector in India's growth, with an emphasis on state-level trends. However, it at least raises questions of regional disparities.

Chadha and Nandwani (2019) study inequalities amongst Indian States. They conclude, thus: "An examination of the performance of development spending (which should mostly benefit the poor) incurred by the states indicates that though faster-growing states showed high spending on the development sector, development spending benefited the rich more effectively than the poor—contrary to the intent behind it—thereby raising inequality in the state".

The literature on special status and the state's revenue is scanty in India. However, some general studies are discussed as under.

"The essence of federalism lies in proper division of powers and functions among various levels of government to ensure adequate resources for their functioning. (p.1)", Nayak and Sathpathy (2017). This paper discusses Federal Finance in India in the context of Discretionary Transfers.

Another paper that has been referred to is Panda (2007). "The paper empirically examines the impacts of federal transfers on States' tax efforts and expenditure taking into consideration a panel data set of 22 Indian States for the time-period 1980-81 to 2007-08 (p.20)", Panda (2017). It discusses the impact of Federal transfers on States' tax efforts. To some extent the conclusions are relevant.

"States' aggregate 'own revenues' (the total of 'own tax revenues' and 'own non-tax revenues'), as a percentage of GDP, showed an upward trend, increasing from 7.0 per cent in 2004-05 to 7.6 per cent in 2012-13, (p.8)", Sucharita (2016). The above paper dwells upon the trend in State's own revenue in recent times but does not talk about the Special States.

"Recognising the fact that the financial resources of the States may prove inadequate for undertaking development activities, the framers of India's constitution have made elaborate arrangements relating to the flow of funds from the Centre to the States, (p.64)", Umesh (2015).

2.1 Rationale

The ranking, in Table 1, clearly shows the advantageous position in which non-Special states are placed. All the top states in the country are from the non-special states. This ranking, may however, conceal the growth pattern of GSDP of special states. Also,

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

there is a special interest in the growth pattern of Jammu and Kashmir, whose special status has been undone recently (5th August 2019). It was stated that Art. 35 and Art. 370 were a hinderance to the growth of Jammu and Kashmir state.

The above papers, hence, provide the ignition for our study. The present study looks into such regional and state-wise differentials in growth patterns, especially in the light of special status accorded to different deviant states. There is a strong underlying belief that Fiscal stimuli not only have an overall growth impact but also may result in either encouraging such special states or isolating them. The main economic argument to remove Special Status from J&K has been to 'integrate' it into 'rest of India'. Growth patterns would throw light on the validation of this policy initiative.

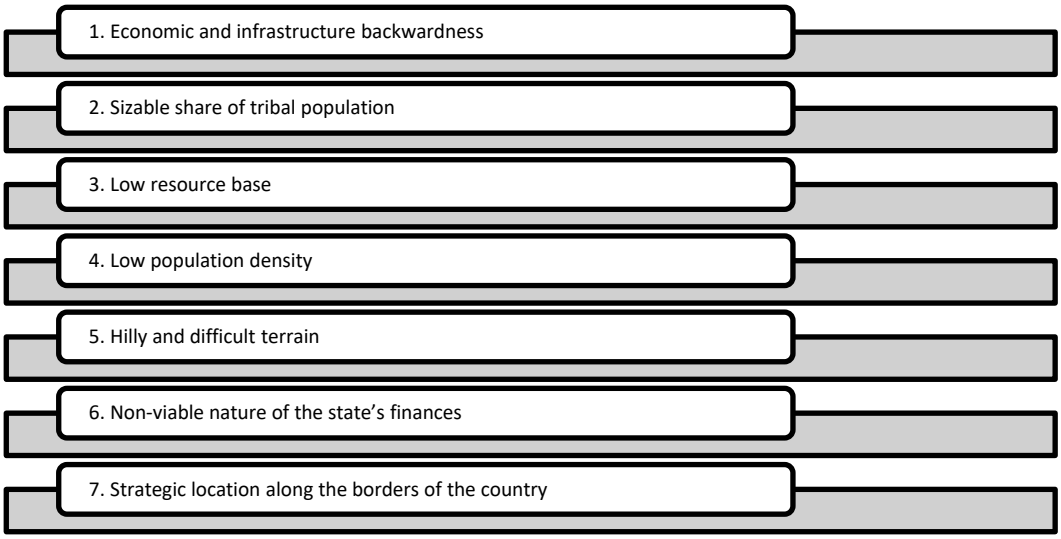
The literature, thus far, has not emphasized the importance of studying growth patterns amongst Indian States, especially from the point of view of the provision of special category status. There is an imperative to do so, especially in the light of the revocation of such status from J&K. There are grave concerns that the fate of other states in India may be decided on similar lines. This paper attempts to provide an objective insight into the past trends in SGDP and aims at throwing light on the very imminent problem.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

2.2 Special category status

Given below in Figure 1 are the seven parameters for special category status to be granted.

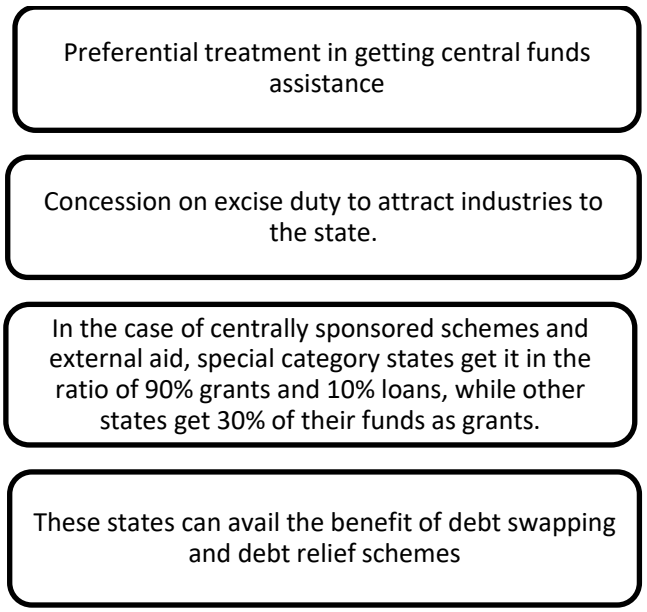
Figure 1: Parameters for Special Category Status



Source: Compiled by author

In Figure 2, mentioned below, there are many 'Benefits for States under SCS'. A significant 30 per cent of the Centre's gross budget goes to the special category states. Thus, the main research question is that if such a large share of the central budget goes to these special states then is it justifies as per their performance in terms of generation of own revenues by such states.

Figure 2: Benefits to Special States



Source: Compiled by author

In 1969, during the 5th finance commission the concept of a special category state was first introduced. The aim was to support disadvantaged states with preferential treatment in the form of central assistance and tax breaks. To begin with only three states Assam, Nagaland and Jammu & Kashmir were granted this status. Ever since the initiation eight more have been included. This makes it a total of 11 Special States. The additional states to be added were - Arunachal Pradesh, Himachal Pradesh, Manipur, Meghalaya, Mizoram, Sikkim, Tripura and Uttarakhand.

Given below is a list of non-Special States. West Bengal has not been included because of the non-availability of data. Telangana has data only from 2005-06. The list of Special states is complete and consists of 11 states as given in Table 2.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Table 2: Non-Special and Special States

Non-Special States	Name	Special States	Name
1	A & N Islands	1	Arunachal Pradesh
2	Andhra Pradesh	2	Assam
3	Bihar	3	Himachal Pradesh
4	Chandigarh	4	Jammu and Kashmir
5	Chhattisgarh	5	Manipur
6	Delhi	6	Meghalaya
7	Goa	7	Mizoram
8	Gujarat	8	Nagaland
9	Haryana	9	Sikkim
10	Jharkhand	10	Tripura
11	Karnataka	11	Uttarakhand
12	Kerala		
13	Madhya Pradesh		
14	Maharashtra		
15	Odisha		
16	Puducherry		
17	Punjab		
18	Rajasthan		
19	Tamil Nadu		
20	Telangana		
21	Uttar Pradesh		

Source: Compiled by author

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

3.0 Data and Methodology

Herein, we discuss the data sources and the methodology used.

3.1 Data

The Gross State Domestic Product is measured in value terms. It is measured in current prices and the unit of measure is Rs. Lakhs. The main data source is from the Central Statistics Office, Ministry of Statistics and Programme Implementation, Government of India (Table 15: Gross State Domestic Product at Factor Cost).

The period of study is essentially from 1993-94 to 2015-16. Technically, the data refers to the financial years 1993-94 to 2015-16. But for empirical purposes, we have taken the year-end as the cut-off. In the case of almost all Non-Special States' data is for the whole period. Consistent data were not available for West Bengal. For the latest state is Telangana whose data is available only from 2014-15 has been eliminated. This leaves a database of 11 Special States and 20 Non-Special States.

3.2 Methodology

A few basic tools of measurement and analysis have been used. Given below is the main research question.

Research Question 1: Since 30 per cent of Central Budget goes to Special States do these states justify such status as per their performance in terms of generation of SGDP growth?

3.3 Regression analysis

Here we have applied semi-log growth equations to measure the instant and annual compound growth rates.

We have regressed the log of each variable with respect the time. Therefore, regression equation can be written as follows in exponential form:

$$Y = e^{\alpha + \beta t} \quad \ldots. (1)$$

Taking log of both sides and adding an error term;

$$\text{Ln } Y = \alpha + \beta t + \mu t \quad \ldots. (2)$$

Where Ln Y = natural log of variable Y

α = intercept term

β = slope of the regression equation

t =time (1993-94 to 2015-16)

μt = error term.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

3.4 Growth analysis

The advantage of a semi-log growth equation is that it directly tells us the growth rate where

β = instantaneous growth rate.

The Annual Compound Growth Rate is derived as:

ACGR= Anti-Log (β) -1

This gives the growth rate in the whole period under consideration while instantaneous growth rate tells us the growth at a point of time. Several comparisons, in respect of ACGR and Instant Growth Rate, are made while analysing growth patterns:

3.5 Convergence index

We now form a relative index of growth that straight away gives the comparative picture of the growth of Non-Special and Special States in respect of own revenue.

Relative Index of Growth = Ri = Rn/Rs

Rn= SDP of Non-Special States

Rs = SDP of Special States

Further, we regress the index over time through a semi-log equation that is taking the log of the relative index as the dependent variable.

$$Ri = e^{\alpha + \beta t} \quad \ldots. (3)$$

Taking log of both sides and adding an error term;

$$Ln\ Ri = \alpha + \beta t + \mu t \quad \ldots. (4)$$

where Ln Ri = natural log of variable Y

α = intercept term

β = slope of the regression equation

t =time (1993-94 to 2015-16)

μt = error term.

Here β gives the proportionate growth rate of Non-Special to the Special States.

Research Question 2: Does the growth rate SGDP of Special States converge with that of Non- Special States?

The non-Special States have a natural advantage of performing better than the Special States because they do not suffer from certain disadvantages as the Special States do. If by getting the necessary support in the form of special status have the Special States been able to pull up their act? If so then the policy support is justified. Else there should be a debate on whether the support should continue. The answer lies in the Convergence Index. If β is <0 then there is convergence. It implies that the special status is justified.

The next research question is:

How have different States' growth patterns performed during the regimes of NDA vs. UPA governments?

To address this question, we estimate a set of Semi-log regression equations. The Semi-log growth equations are loaded with three intercept and three corresponding slope dummies. Through a similar model, we also verify the estimated growth pattern of former J&K State during four political regimes.

Here, we compare the growth rates during four political regimes.

- UPA 1- from 1994 to 2004;
- NDA 1- from 2004 to 2009;
- UPA 2- from 2009 to 2014; and
- NDA 2- from 2014 to 2016.

(We have stopped at 2016 because of the problem of demonetization).

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

This growth analysis enables:

- Comparison between All States, Non-Special Status and Special.
- The pattern of growth in Jammu and Kashmir.

3.6 Graphical analysis

With the help of semi-log equations which are loaded with intercept and slope dummies, we are able to obtain more precise estimates of the predicted level of SGDP. This helps us to compare graphically, the actual and predicted 'Y'. The graphs so obtained clarify whether the equation is a good fit. If the fit is close it shows that the dummy variable exercise is successful in interpreting growth patterns in terms of Political Regimes. While there have been certain historical accidents during these phases the graphs are able to capture if the impact of these events is permanent or temporary. This also allows us to observe if there are any abnormal years where the SDPs are abnormally high or low. If 'Y' and 'Predicted Y' lie close to each other then it implies that such aberrations have failed to sway the general growth pattern.

4.0 Results and Analysis

Returning to the main question about whether the special status given to 11 states in India is justified, a preliminary graphical analysis is desirable. To begin with, we consider the case of Non-Special states. They are 20 in number and are large states.

4.1 Growth patterns

In the following section, we study the broad pattern of growth amongst different categories of states.

For this purpose, four graphs are presented:

- Growth in Special states;
- Growth in Non-Special states;
- Growth in all States combined; and
- Comparative growth in all categories.

The regression output of the semi-log growth equation is given in Table 3.

Growth in Non-Special States started at a very low level. In the UPA1 period, it grew at 13.78 instant growth rates and 14.78 ACGR. During NDA1 the initial level increased substantially. The growth rate during this period was 7.85 instant and 8.17 ACGR. In the next period, UPA2 the initial level fell further but the growth rate in instant terms was 15.28 and the ACGR was 16.58. The best so far. The best was yet to

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

come. The instant rate in NDA2 was 18.61 and ACGR was 20.46% p.a.! Since these states are the majority states they weigh a lot in terms of the overall growth of All States.

Table 3: Growth Pattern of Non-Special States

SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.999692796			
R Square	0.999385686			
Adjusted R Square	0.999099007			
Standard Error	0.026588273			
Observations	23			
ANOVA				
	Df	*SS*	*MS*	*F*
Regression	7	17.25101109	2.464430156	3486.071321
Residual	15	0.010604044	0.000706936	
Total	22	17.26161514		
	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-256.9242445	16.78227252	-15.30926423	1.45485E-10
Year	0.137863508	0.00840795	16.39680393	5.48901E-11
NDA1	118.5475642	21.05879114	5.629362265	4.79635E-05
NDA2	-97.43369385	41.43434574	-2.351520028	0.03278116
TNDA1	-0.059319696	0.010539923	-5.628095643	4.80765E-05
TNDA2	0.048293369	0.020595187	2.344886105	0.03320656
UPA2	-31.4692462	18.14384051	-1.734431372	0.103337263
TUPA2	0.015550244	0.009081624	1.712275692	0.107435184

Source: Compiled by author

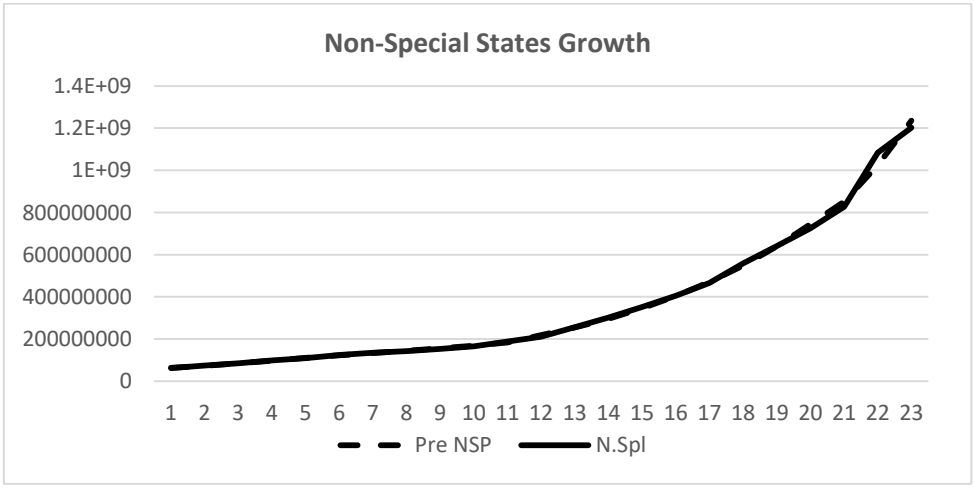
www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

We now examine Figure 4. An exponential growth curve without breaks or jumps is visible. There are no outliers, either in between or at the ends. The entire period, including the Global Financial Crisis, shows that there was no abnormal pattern. The pattern is a smooth irrespective of accounting for four different policy regimes.

The predicted trend closely hugs the actual curve. This speaks for the underlying estimation model. The growth pattern is smooth and exponential. It does not indicate an internal crisis during the whole period. However, this is a single growth trend. We need to study other evidence as well. The pattern in Special states as well as the All India trend needs to be studied (Figure 5).

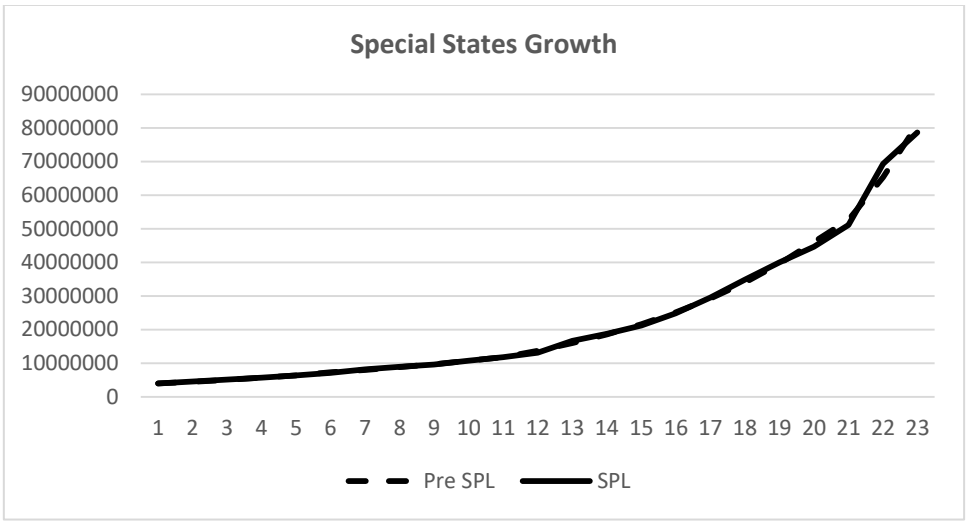
The initial level during UPA1 was quite low (Table 4). The growth rate (instant) was 11.80%. The ACGR was 12.53 p.a. At the beginning of NDA1, the intercept was slightly higher, but it was still low.

Figure 4: Growth in Non-Special States



Source: Compiled by author

Figure 5: Growth in Special States



Source: Compiled by author

The growth rate during this period showed an instant rate of 9.8 and an ACGR of 10.25. There was a slight slippage in growth. In UPA2 period the intercept fell. The growth rate in instant terms was 15.10 and the ACGR was 16.37.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Table 4: Growth Pattern of Special States

SUMMARY OUTPUT					
Regression Statistics					
Multiple R	0.999643				
R Square	0.999286				
Adjusted R Square	0.998952				
Standard Error	0.029281				
Observations	23				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	17.99281	2.570402	2998.00202	1.96875E-22
Residual	15	0.012861	0.000857		
Total	22	18.00567			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	
Intercept	-220.244	18.48184	-11.9168	4.75366E-09	
Year	0.118074	0.009259	12.75179	1.87587E-09	
NDA1	40.93575	23.19145	1.765123	0.097887102	
NDA2	-195.427	45.63047	-4.28282	0.000654028	
TNDA1	-0.02047	0.011607	-1.76381	0.098114612	
TNDA2	0.097145	0.022681	4.2831	0.00065366	
UPA2	-67.3237	19.9813	-3.36934	0.004214789	
TUPA2	0.033547	0.010001	3.354207	0.004347659	

This was a substantial jump. The last period of NDA2 saw a fall in the intercept. However, the growth rate rose to 21.51 in instant terms and an impressive 24.01% per annum!! Clearly, the NDA2 regime had a 'special' for Special States. The graph of Special States also shows a similar pattern of growth. Figure 2 makes it evident that the underlying growth model neatly fulfils the structural pattern of four underlying political regimes. The graph is exponential. It shows undisturbed growth and uniform growth patter irrespective of the political regime. Of course, the post-demonetization period has been excluded from the study. If the latter period had been included maybe the results would have been quite different. The pattern of Special States looks very deceptively close to that of Non-Special states because the difference in the exponential growth patterns is not very large. The actual contribution of Special States seems rather small. The overall growth increases only by 0.32% per annum when we account for how Special States converge to Non-Special States. We now examine all states. Since Non-Special States are the large and dominant states in terms of their size the pattern of Special States and All States are very similar. Given in Table 5 is the regression output of all States in India.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Table 5: Growth Pattern of All States

SUMMARY OUTPUT					
Regression Statistics					
Multiple R	0.999699				
R Square	0.999399				
Adjusted R Square	0.999119				
Standard Error	0.026328				
Observations	23				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	17.29262	2.470374498	3563.847196	5.38909E-23
Residual	15	0.010398	0.000693176		
Total	22	17.30302			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	
Intercept	-254.638	16.61814	-15.32290148	1.43663E-10	
Year	0.136747	0.008326	16.4246702	5.35774E-11	
NDA1	114.1333	20.85284	5.473274669	6.41925E-05	
NDA2	-103.158	41.02912	-2.514262344	0.023825135	
TNDA1	-0.05711	0.010437	-5.471986723	6.43479E-05	
TNDA2	0.051147	0.020394	2.507953347	0.024123976	
UPA2	-33.4782	17.9664	-1.863378102	0.082105875	
TUPA2	0.016559	0.008993	1.841322596	0.085436632	

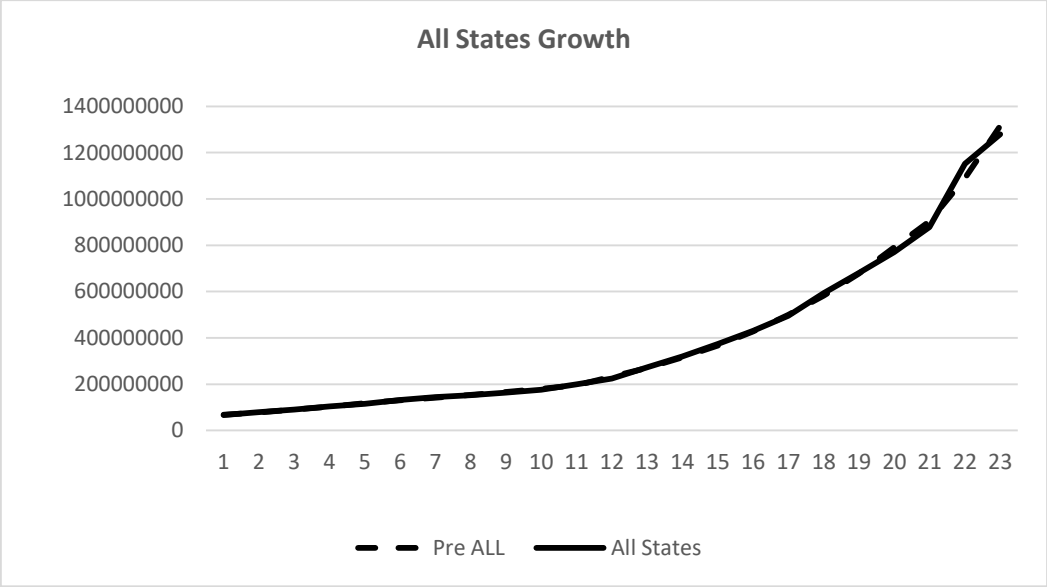
The growth equation of All States shows an instant growth of 13.67 and 14.65 respectively in terms of instant growth and ACGR during UPA1. During NDA1 the ACGR is 8.29% and in UPA2 it goes up to 16.56. Finally, in NDA2 it goes up to 20.61% per annum!! The pattern of All States clearly follows that of Non-Special states.

The graphical analysis is not essentially different from Non-Special states. The dotted line Predicted All States is almost the same as the actual pattern. Once again this speaks for the goodness of fit of our proposed model. It includes the effects of appropriate dummies and a semi-log growth equation which results in a pattern that almost exactly emulates the actual growth pattern on account of the four policy periods. It demonstrates stable growth in all four periods.

In Figures 6 and 7, it is evident that Special states are still at low level in absolute terms. They are backward states and are vulnerable. The level of the dotted line is much lower. But the growth rates of Special States are significantly higher. Although the gap between the two groups may not be very large in terms of growth rates. Also, the overall growth rates are still dominated by the large and Non-Special states.

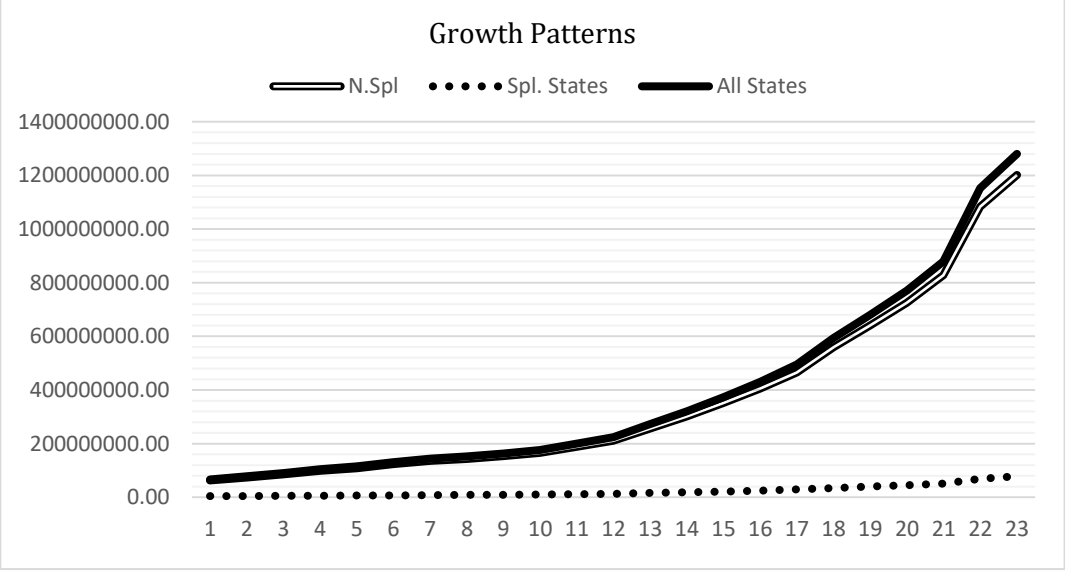
www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Figure 6: Growth in All States



Source: Compiled by author

Figure 7: Comparative Growth Patterns



Source: Compiled by author

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

4.2 Convergence index

Table 6 and Figure 8 explains how the growth pattern of Special states has converged to that of Non-Special states. Although in an overall sense the ratio of Non-Special States' GDP to that of Special States has declined at the rate of almost 0.29% per cent per annum, there are periods of violent swings. These can be explained in terms of political regimes.

Table 6: Convergence Index

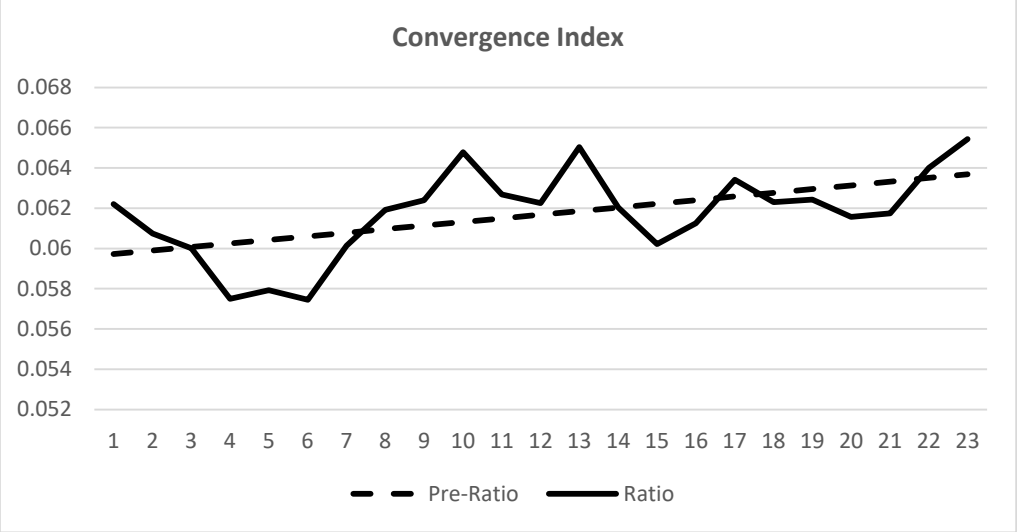
Year	Pre-Ratio	Ratio
1994	0.059727	0.062213
1995	0.059901	0.06075
1996	0.060077	0.06001
1997	0.060252	0.057491
1998	0.060428	0.057927
1999	0.060605	0.057445
2000	0.060782	0.060135
2001	0.06096	0.061916
2002	0.061138	0.062395
2003	0.061317	0.064782
2004	0.061496	0.062679
2005	0.061676	0.062251
2006	0.061856	0.065043
2007	0.062037	0.062037
2008	0.062218	0.06022
2009	0.0624	0.061244
2010	0.062583	0.063407
2011	0.062766	0.062305
2012	0.062949	0.062423
2013	0.063133	0.061564
2014	0.063318	0.061738
2015	0.063503	0.063995
2016	0.063689	0.065433

Source: Compiled by author

The main conclusion of this paper is that Special states, which were disadvantaged have gained during UPA1 and UPA2. They have receded during NDA1 and NDA2 as compared to Non-Special states (Table 7).

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Figure 8: Convergence Index



Source: Compiled by author

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Table 7: Convergence or Divergence

Regime	Growth rate of Ratio of Spl./N.Spl	Convergence/ Divergence
UPA1	(-)1.96%	Convergence
NDA1	(+)1.92%	Divergence
UPA2	(-)0.18%	Convergence
NDA2	(+)2.95%	Divergence

Source: Compiled by author

4.3 Jammu and Kashmir growth

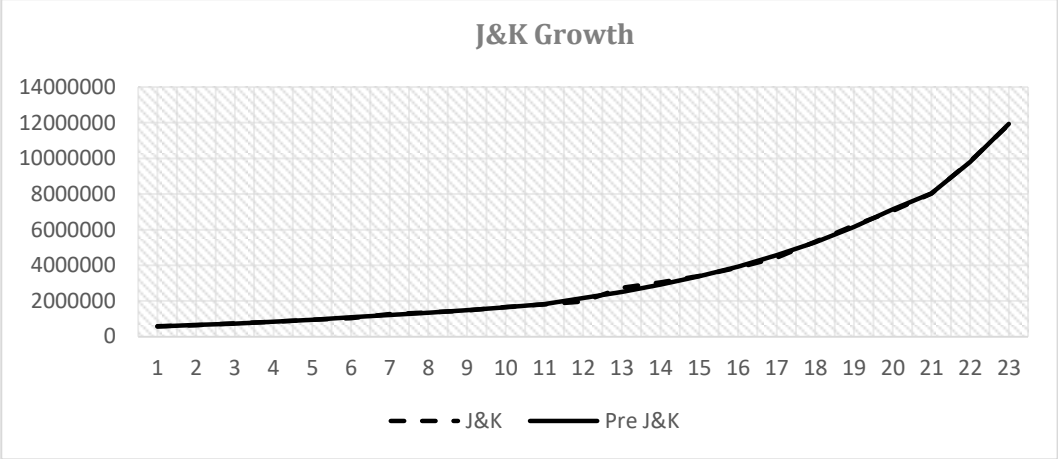
During the UPA1 regime, the instant growth rate was 12.35% p.a. and the ACGR was 13.14% (Table 8). In the beginning of NDA1 the initial level grew but it was not statistically significant. During NDA1 the instantaneous growth fell to 10.31 and the ACGR to 10.86%. However, even this was not significant. This was perhaps because there was an atmosphere of uncertainty. At the beginning of UPA 2 regime the initial level of State GDP fell. But the instantaneous growth rose to 14.90 and the ACGR to 16.06%. Finally, at the beginning of NDA2 regime the intercept certainly fell but the growth rate jumped. During NDA2 the instantaneous growth rose to 19.77 and the ACGR to 21.86!!

Table 8: Jammu and Kashmir Growth

Regression Statistics					
Multiple R	0.999326				
R Square	0.998652				
Adjusted R Square	0.998024				
Standard Error	0.040875				
Observations	23				
ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	18.5721	2.653157	1587.997	2.29819E-20
Residual	15	0.025061	0.001671		
Total	22	18.59716			
	Coefficients	*Standard Error*	*t Stat*	*P-value*	
Intercept	-233.057	25.79986	-9.03326	1.87E-07	
Year	0.123529	0.012926	9.556769	9.04E-08	
NDA1	40.77519	32.37427	1.259494	**0.227098**	
NDA2	-149.209	63.69818	-2.34243	0.033365	
TNDA1	-0.02039	0.016203	-1.25808	**0.227596**	
TNDA2	0.07417	0.031662	2.342585	0.033355	
UPA2	-51.0994	27.89303	-1.83198	0.086884	
TUPA2	0.025471	0.013961	1.824405	0.088073	

The one thing that is apparent is that in spite of different political regimes (Figure 9) the growth pattern of former J&K has been smooth.

Figure 9: J&K Growth Pattern



Source: Compiled by author

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

There are no violent shifts or declines in the growth pattern. That is why the actual and predicted closely follow each other just is the case with all other figures of different growth patterns. This conclusion establishes that J&K has developed resilience in its economy.

4.4 Final note

In the concluding part, we raise another research question for posterity (Table 9). *Research Question 3:* Should Special status continue?

Table 9: Growth in State GDP- 1993-94 to 2015-16

	Instant Growth Rate	Annual Compound Growth Rate	P-Value
All States	13.00%	13.88%	7.6906E-22
Non-Special States	12.98%	13.86%	9.1544E-22
Special States	13.27%	14.20%	9.7890E-31

Source: Compiled by Author

Through an overall comparison, it becomes evident that apparently, the growth rate of Special States is greater. However, if we observe the growth rate of Special States in comparison to the growth of All States then the actual contribution Special States seems rather small. The overall growth increases only by 0.32% per annum. Although it must be said that this is statistically significant. Compounded over 1994-2016 this small contribution would also be somewhat sizeable. In any case, it is an unambiguous conclusion that Special States have performed well. Some have grown faster others are slightly slower. On the whole, their performance outstrips that of Non-Special States.

5.0 Conclusion

While there is a difference amongst Special states as a trend they have done well in terms of economic growth. For the initial 8 years, the performance of Special states was worsening in relation to Non-Special state. Later it started converging at a rate of almost 0.29% per cent per annum. Also, the later growth phase does not belong to any particular political regime. It is partly covered by UPA1, then NDA1 followed by UPA2 and finally, NDA2. This is clearly indicative of a permanent trend.

The main conclusion of this paper is that Special states, that were disadvantaged have gained during UPA1 and UPA2. They have receded during NDA1 and NDA2 as compared to Non-Special states.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

The final question, therefore, is answered in the affirmative. By all counts, Special status should continue. Although, one view is that since the growth of Special states in recent years is irrespective of political regimes the status may be withdrawn because the initial period of backwardness and vulnerability is over.

References

Bhaskar, V. & Gupta, B. (2007). India's development in the era of growth. *Oxford Review of Economic Policy*, *23*(2), 135-142. Retrieved from https://doi.org/10.1093/oxrep/grm021.

Chadha, N. & Nandwani, B. (2019). Growth, development spending, and inequality in Indian states. *Economic and Political Weekly, 54*(11), 45.

Kohli, A. (2006). Politics of economic growth in India, 1980-2005: Part I: The 1980s. *Economic and Political Weekly, 41*(13), 1251-1259. Retrieved from http://www.jstor.org/stable/4418028.

Murthy, K. V. B. (2015). An analysis of socio-economic objectives of India's tax policy. *VISION: Journal of Indian Taxation, 2*(2), 14-27. DOI: 10.17492/vision.v2i2.10178.

Murthy, K. V. B. (2018). State-local financial relations in India. *VISION: Journal of Indian Taxation, 5*(1), 42-59. DOI: 10.17492/vision.v5i1.13275.

Murthy, K. V. B. (2018)a. Trends in own revenue of Indian states: Is special status justified? *VISION: Journal of Indian Taxation, 5*(2), 28-43. DOI: 10.17492/vision.v5i2.14520.

Nayak, C., & Satpathy, P. (2017). Federal finance in India: An analysis of discretionary transfers. *VISION: Journal of Indian Taxation, 4*(1), 1-18. DOI: 10.17492/vision.v4i01.9990.

Panda, P. (2017). Budgetary impacts of central fiscal transfers in India: Evidence from state level Data. *VISION: Journal of Indian Taxation, 4*(2), 20-38. DOI: 10.17492/vision.v4i02.11782.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

Shand, R., & Bhide, S. (2000). Sources of economic growth: regional dimensions of reforms. *Economic and Political Weekly*, *35*(42), 3747-3757. Retrieved from http://www.jstor.org/stable/4409864.

Sucharita, S. (2016). Revenue-expenditure Nexus for India. *VISION: Journal of Indian Taxation, 3*(1), 1-15. DOI: 10.17492/vision.v3i1.7068.

Umesh, M. (2015). Centre-state financial relations: The federal aspects of Indian taxation. *VISION: Journal of Indian Taxation, 2*(1), 64-74. DOI: 10.17492/vision.v2i1.6596.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.100 on dated 4-Dec-2020

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

IEEE Sensors Council

Hardware Emulation of a Biorthogonal Wavelet Transform Based Heart Rate Monitoring Device

Ashish Kumar, *Member, IEEE*, Ramana Ranganatham, Manjeet Kumar, *Member, IEEE* and Rama Komaragiri, *Senior Member, IEEE*

***Abstract*—In this work, a low-power, area-efficient, and high-performance electrocardiogram (ECG) detector that can be used in modern implantable cardiac pacemaker systems is proposed. Qualitative and quantitative analysis indicates that the Biorthogonal-3.1 wavelet transform combined with adaptive slope prediction is best suited to detect various ECG peaks by effectively denoising various artifacts present in an ECG signal. A demand-based wavelet filter bank (WFB) architecture consisting of a series combination of three lowpass filters is the primary reason for the low-power requirements of the circuit. Power consumption and area requirements of the circuit are further reduced by realizing the lowpass filters using lattice wave digital filter realization that reduces the requirement of delay elements, and multipliers count by 80% and 75%, respectively. The proposed ECG denoising scheme using the biorthogonal 3.1 wavelet transform generates a smoother denoised ECG wave by retaining the important morphology of an ECG signal. Adaptive slope prediction criterion-based ECG processing and detection schemes achieve a high detection accuracy of 99.90%, with the lowest error of 0.002%. The proposed scheme is capable of distinguishing between a normal ECG, paced ECG, arrhythmic ECG, and low and high-resolution ECG. The proposed algorithm is then emulated on the Xilinx®-Virtex®-7 FPGA hardware platform. Power consumption, area, delay, and switching energy of the proposed ECG processing scheme are reduced when compared to existing ECG detection schemes and are found to be of 0.493µW, 1.1mm^2, 10ns, and 4.93µJ, respectively.**

***Index Terms*—Electrocardiogram, wavelet transform, wavelet filter bank, ECG processing**

I. Introduction

Cardiovascular diseases (CVD) are one of the primary reasons for the high global mortality rate in modern times. CVDs resulted in the death of 17.5 million people, approximately 31% of the overall global deaths [1]. Medical practitioners are widely open in accepting the usage of wearable ECG devices for biosignal acquisitions, as a better patient examining and supervising device in real-time. Medical institutions require automated equipment to improve cardiovascular analysis. Holter devices are incapable of providing real-time diagnosis and information of arrhythmia, thus limiting their usage in critical conditions. Hence, there is an immediate requirement to develop energy-efficient ECG detection algorithms for implantable and portable cardiac devices. A Block diagram representation of a cardiac device inclusive of an ECG detector is depicted in Fig. 1. Periodic and administered electric stimulus made the implantable device a reality in today's globally emerging world of portable devices. The present implantable and portable cardiac pacemakers are expensive, so not within reach of many. Thus, to become a real lifesaver, portable, and implantable cardiac pacemaker needs to be cheap and yet reliable. The portable pacemaker devices require higher circuit densities and technological breakthroughs so that all the lacunas like single-chamber, asynchronous, non-programmable pacing can be left behind. Technology adds more flavor and variations in the device with benefits like dual-chamber multi-programmability for enhanced efficiency and reliability. Such newer and advanced devices have amplified the storage of important medical data, which may enable early detection, timely diagnosis, and treatment of CVD [2]. With the current progress, the contemporary pacing system has three integral components: a pulse generator, lead, and ECG detector as part of its device architecture.

Amongst the commonly used clinical cardiac tests, ECG analysis provides valuable information about the structure and functionality of the heart. The ECG signal is utilized to study the effect of various drugs on the human heart. The data extracted from ECG is also used to diagnose various ailments like myocardial ischemia, atrial abnormalities, by extracting the three primary parameters in an ECG signal that are P-wave, QRS complex, and T-wave. These ECG features typically have a duration of 0.08-0.10 seconds for P-wave, 0.06-0.10 seconds for QRS complex, and 0.10-0.25 seconds for the T-wave. During ECG signal acquisition, some phenomena like electrode misplacement, high-frequency electromyographic noise, motion artifacts, baseline wander, powerline interference, and muscle contraction deteriorate the signal quality. Among the noise sources mentioned above, motion artifact and muscle contraction deteriorate are the primary sources to deteriorate the ECG signal quality and make the visual and automated diagnosis unreliable.

Various literature discusses power-efficient ECG signal detection techniques. An extensive review of the ECG detection

Submitted on 26 AUG 2020.

Ashish Kumar is with the School of Electronics Engineering, Vellore Institute of Technology, Chennai, 600127 India (e-mail: akumar.1june@gmail.com).

Ramana Ranganatham was with the Department of Electronics and Communication Engineering, Bennett University, Greater Noida, 201310, India (e-mail: ramaragam@gmail.com).

Manjeet Kumar is with Department of Electronics and Communication Engineering, Delhi Technological University, Bawana, Delhi, 110042 India. (e-mail: manjeetchhillar@gmail.com).

Rama Komaragiri is with Department of Electronics and Communication Engineering, Bennett University, Greater Noida, 201310, India (e-mail: rama.komaragiri@gmail.com).

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

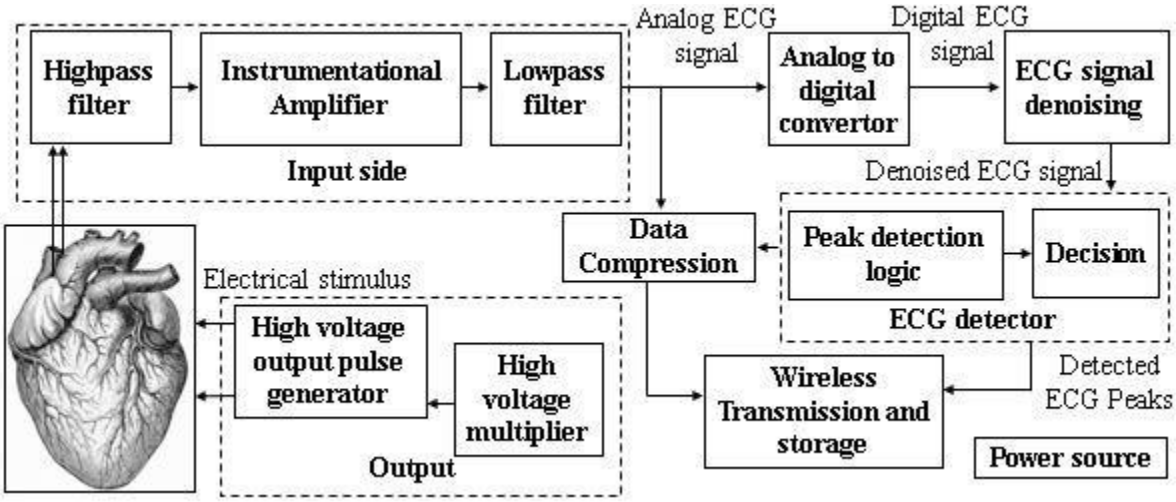


Fig. 1. Implantable/portable cardiac device

techniques presented to date is available in [3-6]. Different research groups introduced various ECG signal detection techniques based on time-domain [7-9], ECG morphology (neural network) [10-12], Time-frequency (wavelet transform, Hilbert transform, empirical mode decomposition (EMD) [13-17], and some combined algorithms (Bandpass filtering and wavelet transform) [18]. An overview of the existing ECG signal detection techniques is shown in Fig. 2. In the present work, ECG processing schemes are evaluated based on robustness to noise, numerical efficiency, and parameter choice. The first real-time ECG detection scheme presented by J. Pan [19] was based on the time-domain analysis. In this technique, high order bandpass filter and thresholding technique are utilized to detect various waves of an ECG signal.

The difficulties encountered with the time domain-based ECG signal detection techniques implemented using finite impulse response filters are computational delay and ringing effects due to the long impulse response. An infinite impulse response filter causes non-linear phase distortion [20]. Zero-phase bidirectional filters [21] overcome the difficulties introduced by finite and infinite impulse filters. Adaptive filtering can attenuate the time-varying noises and retain the ECG signal features of the same frequency range [22-24], which makes bandwidth tuning a difficult task. Adaptive filtering schemes require a pre-defined model of noise to eliminate denoise the signal, thus increasing the computational complexity of the algorithm. The increase in computational complexity degrades the convergence and increases the computational time.

Further, ECG morphology and combined ECG processing schemes have been presented to achieve high peak detection accuracy. Implementing these schemes on the integrated circuit is too complicated. Time-frequency analysis techniques, namely, Hilbert transform, empirical mode decomposition and its modified versions, wavelet transform, have also been utilized for ECG analysis. Time-frequency based approaches provide a very high ECG peak detection rate at the expense of computational complexity. The computational complexity requires a large memory, thus limiting their utility in real-time scenarios [8]. EMD is a specialized analysis which is generally most useful in exploratory analyses [25]. The main advantage of EMD is its ability to estimate subtle changes in time-varying frequency. EMD is not only sensitive to noise but also limits the frequency reconstruction of the signal. In EMD, the reconstructed frequency of each stage is about half that of the previous stage, thus limiting the hypothesis-driven analyses [25]. Considering the tradeoff between robustness to noise, numerical efficiency, and parameter choice, wavelet transform based ECG processing, and detection techniques are found to be most suitable and hence, used in the present work.

In the proposed work, a biorthogonal wavelet transform based modified wavelet filter bank (WFB) [26] to suppress the various low and high-frequency noises present in the ECG signal is implemented on Field programmable gate array (FPGA). Quantitative Measures, namely, inequality, Shannon entropy, uncertainty, cross-correlation, relative entropy, mutual information, and distribution error, are considered to select the optimal base wavelet. Steps involving in wavelet denoising are as follows:

- The approximation and the detail coefficients are obtained using multi-level wavelet decomposition.
- ECG signal was split into a lowpass sub-band (approximation level) using a discrete wavelet transform. Further decomposing the approximation sub-band at multiple scales results in a fine-scale analysis.
- A suitable thresholding technique is selected by Analyzing the detailed coefficients to reconstruct the ECG signal.

After denoising the ECG signal using biorthogonal wavelet transform, an adaptive slope prediction criterion is utilized to check the location of different peaks present in the ECG signal. The proposed work is implemented using System Verilog HDL and synthesized in FPGA on the Virtex®-7 platform.

The rest of the paper is organized as follows: a mathematical description of the wavelet transform is discussed in Section II. The proposed scheme utilized to find the location of different peaks of an ECG signal is discussed in Section III. Experimental results, comparison of the proposed work with the existing literature, respectively, are discussed in Section IV. Finally, Section V concludes the paper.

II. Proposed Work

Various R-peak detection algorithms have been proposed and recommended for removal of such noises and distractions from the data for better and effective detection of the peaks in ECG signals. The previously designed wavelet transform based R-

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

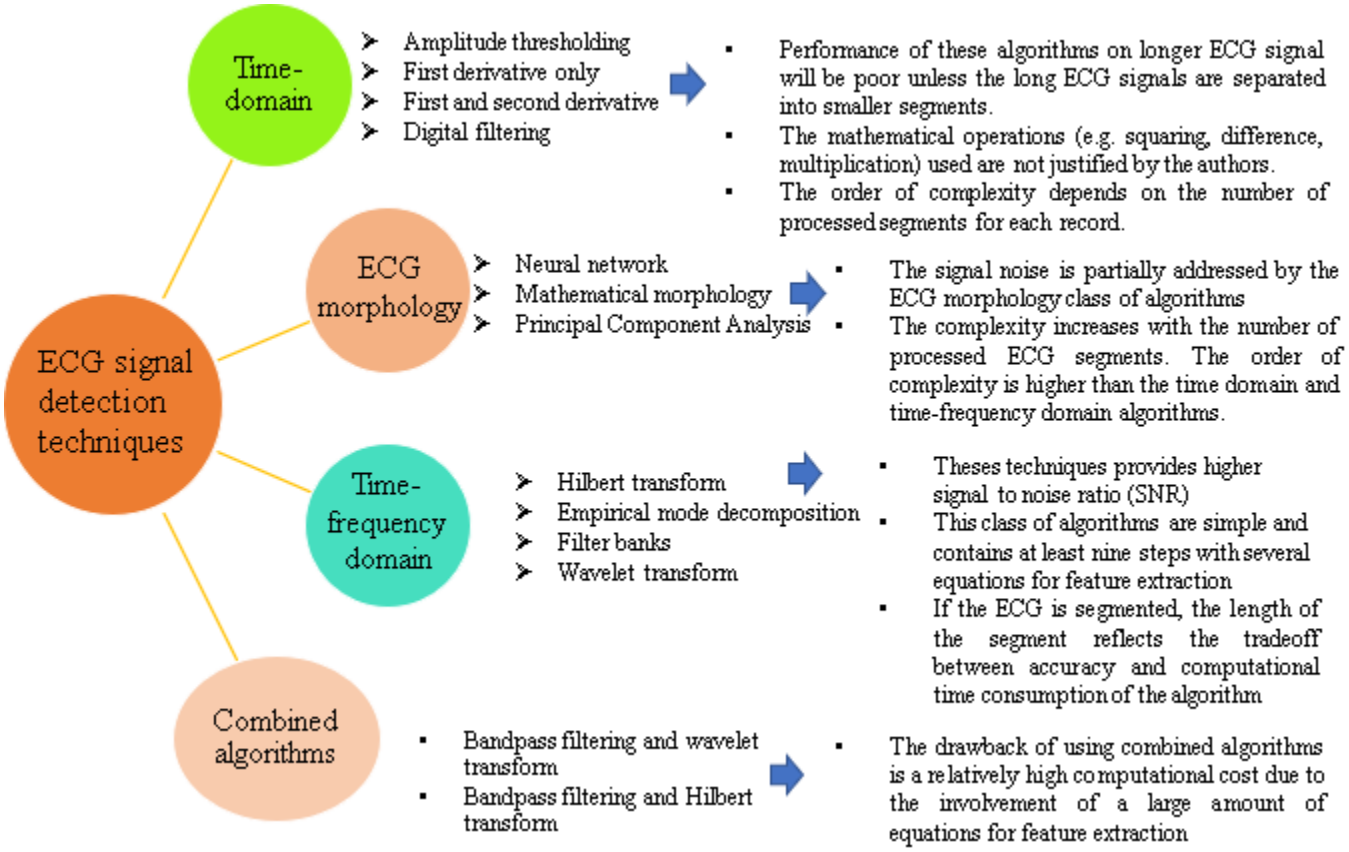


Fig. 2. Overview of the existing ECG signal detection techniques

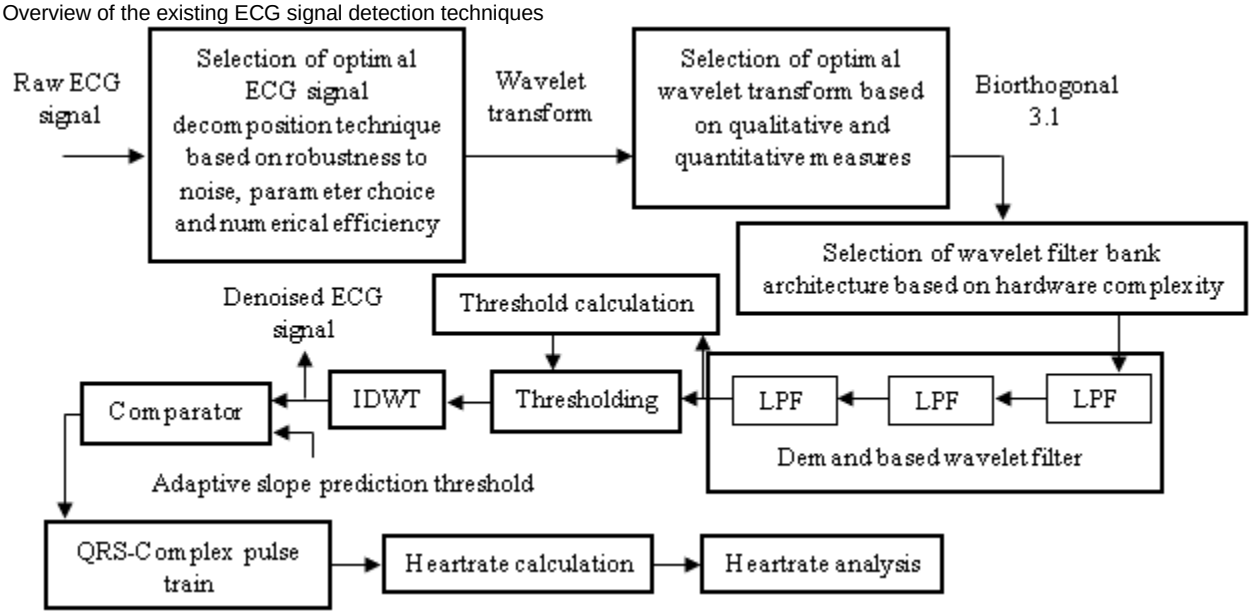


Fig. 3. Block diagram of the proposed QRS-complex detector

peak detection schemes [26-29, 32] utilize the basic signal decomposition process of Mallat's multiresolution analysis, which splits the ECG signal into a lowpass sub-band (approximation level) and highpass sub-band (detail level). The use of lowpass and highpass sub-band requires a large number of adders, multipliers, and delay elements and hence, increases the hardware cost. To minimize the hardware cost, a demand-based biorthogonal wavelet filter bank is used in the proposed scheme. Hardware complexity of the proposed scheme is further reduced by realizing the biorthogonal wavelet transform based lowpass filter or approximation filter using lattice wave digital filter (LWDF) realization. An adaptive slope prediction threshold is also used to increase the peak detection performance. A block diagram representation of the proposed scheme ECG detection scheme is shown in Fig. 3 and is explained below.

The input ECG signals are recorded using iworx® IX-TA-220 recorder [30]. The ECG data is then digitized at 360 Hz sampling frequency with an 11-bit resolution over a 10-mV range. Various other ECG databases available on *physionet.org* [31] are also used to evaluate the performance of the proposed scheme. *Physionet.org* provides a collection of physiological

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

signals which are free for research. *Physionet.org* has a verity of ECG data [33] that includes ECG database contaminated with motion artifacts, MIT-BIH arrhythmia database, fantasia database, QT database, PTB database, to name a few. Based on different physiological conditions, input ECG data are then classified into normal ECG signal, ECG signal with right bundle branch block, ECG signal with left bundle branch block, ECG signal with premature atrial contraction, and ECG signal with premature ventricular contraction [26]. Different white Gaussian noise signals are generated and added to the input ECG signal. Adding different Gaussian noise signals alters the input signal to noise ratio (SNR) of the ECG signal, thus allowing to test the proposed scheme for critical cases. Noisy ECG signals are then decomposed using a biorthogonal wavelet filter bank to suppress the noises. The procedure for selecting wavelet and wavelet filter bank architecture is discussed below.

A. Mathematical Background of Wavelet Transform

Define Wavelets are very useful in analyzing those signals composed of slowly varying components interspersed with abrupt changes. Wavelet transforms are classified as continuous wavelet transform (CWT) and discrete wavelet transform (DWT).

The mathematical background of DWT is as follows [26]. A discrete-time signal *x(n)* is decomposed using wavelet transform as follows:

$$W_{\phi}(p_0,q)=\frac{1}{\sqrt{M}}\sum_{n}x(n)\phi_{p_0,q}(n) \tag{1}$$

$$W_{\psi}(p,q)=\frac{1}{\sqrt{M}}\sum_{n}x(n)\psi_{p,q}(n),\quad p\geq p_0 \tag{2}$$

Here, '*M*' is a normalizing term. $W_{\phi}(p_0,q)$ and $W_{\Psi}(p,q)$ are the scaling and wavelet functions, respectively. $\psi_{p,q}$ and $\phi_{P_0,q}$ represent the transformation kernels. Normally p_0 is chosen to be zero. The above equations can be rewritten as the following.

$$W_{\psi}(p,q)=\sum_{m}h_{\psi}(m-2q)W_{\phi}(p+1,m) \tag{3}$$

Similarly, the scaling function can be written as in Eq. (4).

$$W_{\phi}(p_0,q)=\sum_{m}h_{\phi}(m-2q)W_{\phi}(p_0+1,m) \tag{4}$$

The decomposed discrete-time signal is reconstructed using the inverse wavelet transform. Inverse wavelet transform can be computed using Eq. (5)

$$x(n)=\frac{1}{\sqrt{M}}\sum_{q}W_{\phi}(p_0,q)\phi_{p_0,q}(n)+\sum_{p=p_0}\sum_{q}W_{\psi}(p,q)\psi_{p,q}(n) \tag{5}$$

B. Selection of Wavelet Transform

Use Wavelets developed in the recent past have paved the way for applying a wavelet transform to biomedical signal analysis. Existing literature showcases thirteen wavelet families well archived and cataloged. The study so far in the literature put forward a clear perspective for selecting the best-suited base wavelet for functional signal processing. The fundamental properties support only the qualitative acceptance and fitment for a specific application, thus making a study of quantitative measures in the selection of base wavelet a much-required approach. The present visual differentiation and contrast studies are not sufficient to accurately match the shape of a signal to that of a base wavelet transform, thus establishing an accurate quantitative matching is much required. Thus, the selection criteria of a wavelet suitable for ECG signal analysis made by considering important properties like symmetry, regularity, orthogonality, compact support, a support width, filter length, shape of the wavelet transform, and the number of vanishing moments. Based on properties, the classification of wavelet transforms listed in Table I. Here, N_d and N_r are decomposition order and reconstruction order, respectively. From Table I, it is clear that the biorthogonal wavelet transform satisfies most of the properties required and thus selected for analyzing ECG signals.

Mathematically, energy, Shannon entropy, cross-correlation, minimum description length, mutual information, and relative entropy are the quantitative measures used to evaluate the performance of base wavelet for ECG signal denoising [32].

1) Energy and Shannon Entropy

Features of a signal are characterized using the energy content of a signal. Eq (6) gives the energy content of a signal *x(n)* (*n = 1, 2, 3…., Z*) calculated using its wavelet coefficients. The energy associated with each scaling parameter '*s*' is calculated using Eq. (7).

$$energy=\sum_{s}\sum_{n}|c(s,n)|^2 \tag{6}$$

$$energy(s)=\sum_{n=1}^{Z}|c(s,n)|^2 \tag{7}$$

Here, '*Z*' is equal to the number of wavelet coefficients, and *c(s,n)* represents wavelet coefficients. Shannon entropy (sh_e) is a measure of uncertainty associated with random variables. Hence, sh_e helps to describe the energy distribution of wavelet coefficients, which is calculated using Eq. (8).

$$sh_e=-\sum_{n=1}^{Z}P_n\cdot\log_2(P_n) \tag{8}$$

Here, P_n stands for the energy probability distribution of wavelet coefficients and is calculated using Eq. (9).

$$P_n=\frac{|c(s,n)|^2}{energy(s)} \tag{9}$$

With $\sum_{n=1}^{2}P_n=1$, and if $P_n=0$, then $P_n log_2 P_n=0$. Entropy (sh_e) of the wavelet coefficients is bounded by $0\leq entropy(s)\leq log_2 Z$. The base wavelet that satisfies the condition of maximum energy, minimum sh_e should be the appropriate wavelet transform for ECG signal denoising.

1) Mutual Information and Relative Entropy

The amount of information that signal *S* contains about signal *X* is measured using mutual information (I_{mutual}) and is

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

TABLE I
CLASSIFICATION OF WAVELET WITH THEIR MAIN PROPERTIES

Wavelet family	Orthogonality	Compact support	Support width	Regularity	Symmetry	Filter length	Number of vanishing moments
Haar	Orthogonal	Yes	1	No	Yes	2	1
Daubechies	Orthogonal	Yes	2N-1	No	No	2N	N
Coiflets	Orthogonal	Yes	6N-1	No	No	6N	2N-1
Symlets	Orthogonal	Yes	2N-1	No	No	2N	N
Biorthogonal / reverse biorthogonal	Orthogonal (biorthogonal)	Yes	$2N_r+1$, $2N_d+1$	Yes	Yes	$Max(2N_r,2N_d)+2$	N_r
Discrete Meyer wavelet	Not Orthogonal	No	NA	No	Yes	NA*	NA

calculated using Eq. (10).

$$I_{mutual} = \left(\sum_{s\in S}\sum_{x\in X} p(s,x)\log p(s,x) - \sum_{s\in S}\sum_{x\in X} p(s,x)\log[p(s)p(x)] \right) \quad (10)$$

The average mutual information between two signals *S* and *X* can be calculated by subtracting the entropy of pair from the sum of two self-entropies. The amount of shared information contained by signals *S* and *X,* respectively, in the wavelet thresholding, is represented by mutual information.

The distance between the probability distributions of two signals *S* and *X* is measured using relative entropy ($E_{relative}$) and calculated using Eq. (11). Relative entropy is zero only when the probability distributions of both the signals are equal.

$$E_{relative} = D(S \| X) = \sum_{s\in S} p(s)\log\frac{p(s)}{p(x)} \quad (11)$$

2) Cross-Correlation

Cross-correlation (γ) between the ECG signal and wavelet filter is computed using Eq. (12).

$$\gamma = \frac{\sum_{n}\left[\left(x(n)-m_x\right)*\left(w(n)-m_w\right)\right]}{\sqrt{\sum_{n}\left(x(n)-m_x\right)^2}\sqrt{\sum_{n}\left(w(n)-m_w\right)^2}} \quad (12)$$

Here, *x(n)* represents the ECG data set, *w(n)* represents wavelet data set, m_s is mean of ECG data, and m_x is mean of wavelet data. Base wavelet with maximum cross-correlation coefficients will be selected for the proposed work. Finally, the Optimal base wavelet is the one which provides the shortest description of ECG data and base wavelet. Based on the qualitative measures, biorthogonal 3.1 wavelet transform is found to be best amongst other existing wavelet families, namely, Haar, Daubechies, Coiflets, symlets, discrete Meyer wavelet to isolate the QRS complex and to reduce high-frequency noise in the ECG signal. The present visual differentiation and contrast studies only provide qualitative measures to match the shape of an ECG signal to that of a base wavelet transform. Establishing a qualitative match of the shape of an ECG signal and that of a wavelet transform is required. Table II compares the performance of various wavelet transforms using quantitative measures. The quantitative measures used in this study are the minimum Shannon entropy, the highest energy-to-entropy ratio, large mutual information, high relative entropy, and high normalized correlation coefficient.

C. Selection of Wavelet Filter Bank Architecture

Various wavelet filter bank architectures for different signal processing applications have been presented in the existing literature. For a given precision, VLSI implementation of a wavelet filter bank architectures requires minimization of area, power, and memory. The performance of a conventional wavelet filter bank can be improved by using undecimator and decimator-based wavelet filter banks [34,35]. Although the use of undecimated wavelet filter bank architecture provides the benefit of translation-invariance, the requirement of a large number of registers and a constant clock is its major drawback. The limitations posed by an undecimator wavelet filter bank architecture are overcome by using decimator wavelet filter

TABLE II
PERFORMANCE OF DIFFERENT WAVELET TRANSFORM ON QUANTITATIVE MEASURES

Wavelet Transform	Energy (J)	Shannon entropy	Energy to Shannon entropy ratio	Mutual information	Relative entropy	Normalized correlation coefficient
Haar	35.76	5.22	6.850575	0.86	0.69	0.15
Db15	62.82	3.92	16.02551	0.99	1.05	0.25
Db20	64.61	3.55	18.2	1.31	1.16	0.61
Sym17	63.51	4.54	13.98899	0.95	0.12	0.50
Sym20	64.57	4.51	14.31707	1.19	0.19	0.42
Coif1	36.95	3.12	11.84295	0.83	0.28	0.18
Coif4	58.51	3.17	18.45741	1.32	0.17	0.37
Coif5	62.67	3.06	20.48039	1.46	1.14	0.87
Bior2.6	56.93	5.73	9.935428	1.53	0.15	0.23
Bior2.8	59.17	5.84	10.13185	1.50	0.81	0.67
Bior3.1	**67.73**	**2.92**	**23.19521**	**1.76**	**1.93**	**1.00**
Bior3.7	63.16	4.53	13.9426	0.95	0.19	0.69
Bior6.8	62.46	4.59	13.60784	1.68	1.59	0.51
RBior3.1	66.23	3.28	20.19207	1.70	1.79	0.93
RBior3.7	64.05	5.00	12.81	0.87	0.81	0.76
RBior6.8	62.98	4.97	12.67203	1.59	1.22	0.71

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

bank architecture. However, the use of the parallel combination of lowpass and high pass filters to decompose signals at each level of decomposition requires nearly the same amount of hardware. Hence, a new wavelet filter bank architecture to suppress different noises present in the ECG signal is required. A demand-based wavelet filter bank architecture that decomposes the noisy ECG signal using only three lowpass filters is proposed in this present work. ECG signal '*x(n)*' is first filtered with a special lowpass filter to yield lowpass sub-bands. Half of the samples are discarded after filtering as per the Nyquist criterion. A filter that typically has a small number of coefficients results in a better computational performance. The filter can reconstruct the sub-bands while canceling any aliasing that occurs due to down-sampling. The process repeats itself for the next two decomposition levels. Reason for modifying the wavelet filter bank architecture and selecting the third level of wavelet decomposition is as follows: As per the existing literature, the ECG signal has a frequency range of 0.05-150 Hz and out of which QRS-complex (R-peak) has a frequency range of 3-40 Hz [36]. The proposed wavelet filter bank provides a frequency range of 0-45 Hz, which is similar to the frequency range of the QRS complex, as shown in Fig. 4. Also, the proposed wavelet filter bank architecture decreases the hardware complexity by reducing the High pass filter's count from four to zero.

D. Processing ECG signal

Continuous efforts are made by various researchers to improve the detection capabilities of the vital features of an ECG signal. The use of thresholding approaches for ECG peak detection is preferred due to the following reasons: simple, numerically efficient for the detection of different peaks present in an ECG, minimum memory storage, and provide a high detection accuracy. The present work utilizes an adaptive slope prediction criterion to check the location of different peaks present in the

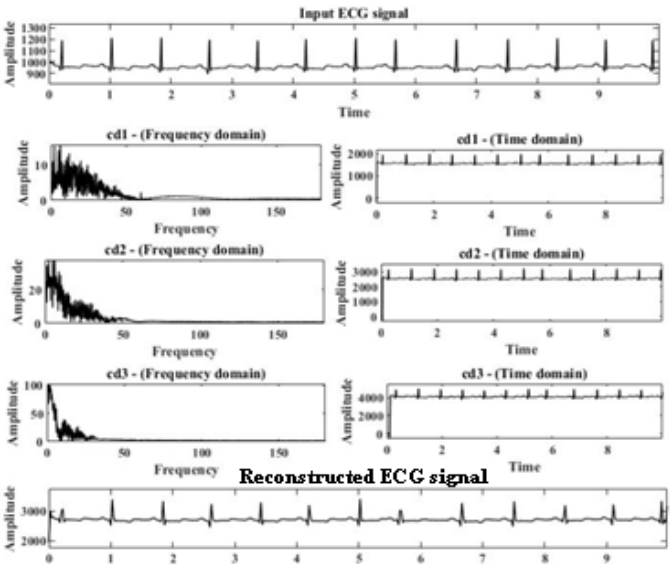


Fig. 4: The frequency response of ECG signal for different wavelet decompositions. 1) The plot on the top shows the input ECG signal (10 seconds of 100m.mat), 2) The plot on the second row shows the first wavelet decomposition cd1 and its frequency components, 3) The middle (third row) plots show the second wavelet decomposition cd2 and its frequency components, 4) The fourth-row plots show the third wavelet decomposition cd3 and its frequency components, 5) The bottom plot shows the denoised ECG signal using the proposed algorithm

ECG signal. A similar slope prediction criterion to detect different peaks present in an ECG is used in [26], but few changes have been made to improve the peak detection accuracy. If the proposed scheme enables us to find an R-peak in 1000ms, that means the interval between two consecutive R-peaks is too large. A large R-R peak interval represents irregularity in the heart's functionality. Hence, the entire region of that time interval is presented for diagnosis to the physician or cardiologist.

III. Simulation and Results

The effectiveness of the biorthogonal 3.1 wavelet transform and adaptive slope prediction criterion-based ECG detection scheme is verified by evaluating the performance on different ECG signals. ECG signals of different duration, ECG signals of a different physiological condition are considered in the evaluation. Firstly, the optimal wavelet transform is selected based on qualitative and quantitative measures. Then, wavelet decomposition level and wavelet filter bank architecture are selected to denoise the various artifacts present in an ECG. The metrics under consideration are hardware complexity, denoising capabilities, power consumption, and area. After removing the various noises, an adaptive slope prediction criterion is utilized to detect the location of different peaks of an ECG signal. Finally, the performance of the proposed scheme is compared with the existing literature. For a fair comparison, the performance of the existing ECG detection schemes is calculated using a similar environment.

A. ECG Signal Denoising

The proposed work utilizes biorthogonal 3.1 wavelets transform to remove the various artifacts present in an ECG. The optimal selection of base wavelet is made using qualitative and quantitative measures. The denoising capabilities of the proposed demand-based biorthogonal 3.1 wavelet transform based ECG denoising scheme is evaluated using the signal-to-noise ratio (SNR), percent root-mean-square difference (PRD), and mean square error (MSE). Each of these metrics provides different qualitative and quantitative facets of the denoised signals with the help of the clean and noisy ECG. Ideally, high SNR and low PRD/MSE values signify better denoising capabilities.

The results of the proposed method have been compared with the four existing wavelet transform based ECG denoising methods [34, 35, 37, 38]. The denoising results of the techniques, as mentioned earlier, are compared for two separate ECG records from motion artifact contaminated ECG database at four different input SNR noise levels. Through close inspection, one can see that the proposed method generates smoother denoised results and retains the essential morphologies of the ECG compared to existing approaches.

Denoising performance of the various ECG denoising methods and their comparison with the proposed method for different input SNR is shown in Table III. A close inspection of Table III shows the proposed method better when compared to the other ECG denoising methods. The proposed ECG

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

denoising scheme achieves the highest SNR of 54.92 dB and lowest PRD(%)/MSE of 11.41 % and 0.008, respectively.

B. Processing ECG Signal

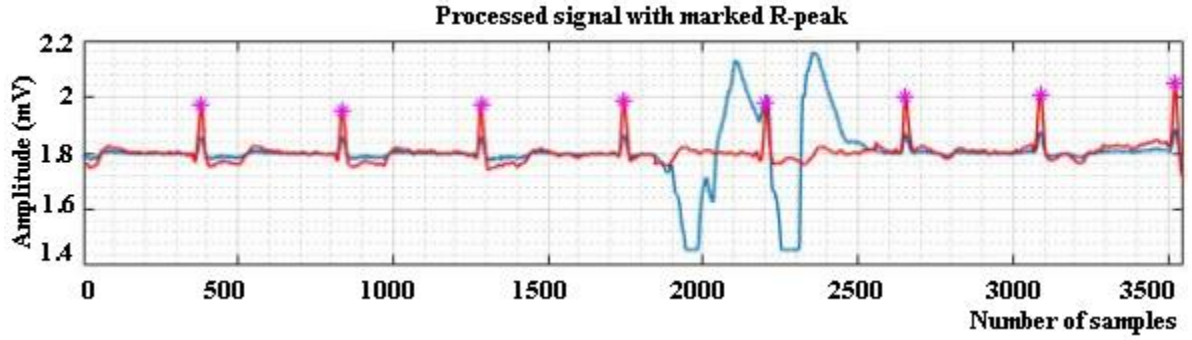
Detection accuracy (DAcc), sensitivity (Se), detection error (DErr), and time consumption (TC) are the metrics used to evaluate the performance of the adaptive slope predication criterion-based ECG detector. The performance of the proposed adaptive slope prediction-based ECG detector for different ECG databases is shown in Table IV. From Table IV, it has been observed that the proposed ECG detector achieves the highest average detection accuracy, average sensitivity with a lower average detection error, and time consumption. The essential ECG components detected by the algorithm is shown in Fig. 5. Fig. 5 (a) shows the output of the muscle contraction noise corrupted ECG signal, whereas Fig. 5 (b) shows the output of motion artifact noise corrupted ECG signal. As discussed earlier that the muscle contraction and motion artifact are the prominent noises that deteriorate the signal quality, hence the performance of the proposed scheme needs to evaluate these two types of signals. From Fig. 5, it is visible that the proposed scheme outperforms on muscle contraction and motion artifact noise corrupted ECG signals. Performance comparison between the existing methods and the proposed method is discussed in Table V. A close inspection of Table V shows that the proposed ECG detector performs better as compared to the existing ECG detectors. The proposed ECG detector can differentiate between normal ECG, paced ECG, arrhythmic ECG, low-quality ECG, and high-quality ECG with higher detection accuracy ranging from 99.86% to 99.93%.

C. FPGA Implementation

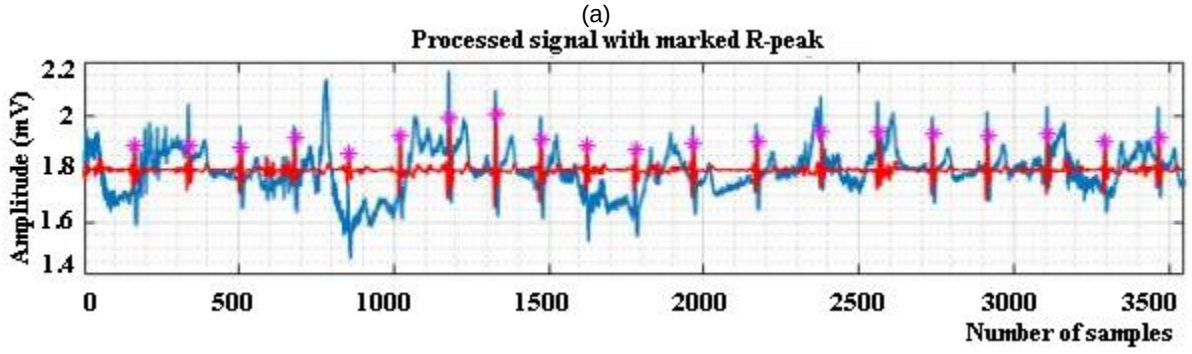
The proposed ECG detector using biorthogonal 3.1 wavelet transform and an adaptive slope prediction technique have four major blocks, namely, wavelet filter bank, thresholding, comparator, and a counter. System Verilog hardware description language and Xilinx Vivado® design suit are used to implement the proposed scheme and complete its functional verification. Area, power, and delay requirements of the proposed scheme are calculated by implementing the detector on the Xilinx Virtex-7 FPGA. The top-level RTL view of the proposed wavelet filter bank architecture and complete

TABLE III
DENOISING PERFORMANCE OF THE VARIOUS ECG DENOISING METHODS FOR DIFFERENT INPUT SNR

Reference		[34]	[35]	[37]	[38]	Proposed
Input SNR (-10 dB)	Average *SNR* (dB)	27.702	27.974	28.256	27.942	**37.180**
	Average *PRD* (%)	22.19	21.71	24.11	23.85	**12.08**
	Average *MSE*	0.120	0.113	0.092	0.102	**0.036**
Input SNR (-5 dB)	Average *SNR* (dB)	28.111	29.108	30.280	29.353	**40.964**
	Average *PRD* (%)	18.38	19.02	18.09	17.52	**11.94**
	Average *MSE*	0.100	0.098	0.089	0.103	**0.014**
Input SNR (5 dB)	Average *SNR* (dB)	30.372	31.979	32.120	32.312	**46.829**
	Average *PRD* (%)	16.27	15.93	13.99	13.44	**11.87**
	Average *MSE*	0.052	0.065	0.048	0.075	**0.008**
Input SNR (10 dB)	Average *SNR* (dB)	32.516	33.052	33.989	33.693	**54.630**
	Average *PRD* (%)	13.59	13.41	13.04	12.99	**11.41**
	Average *MSE*	0.011	0.027	0.021	0.041	**0.008**



(a)



(b)

Fig. 5. R-Peak detection performance of the proposed technique on (a) muscle contracted ECG signal (b) motion artifacted ECG signal

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

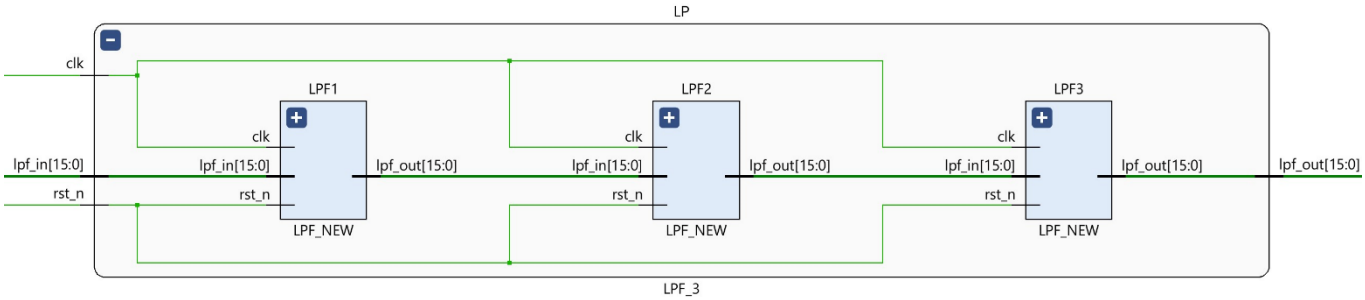
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

TABLE IV
PERFORMANCE COMPARISON OF THE PROPOSED SCHEME FOR DIFFERENT CLASS OF ECG DATA

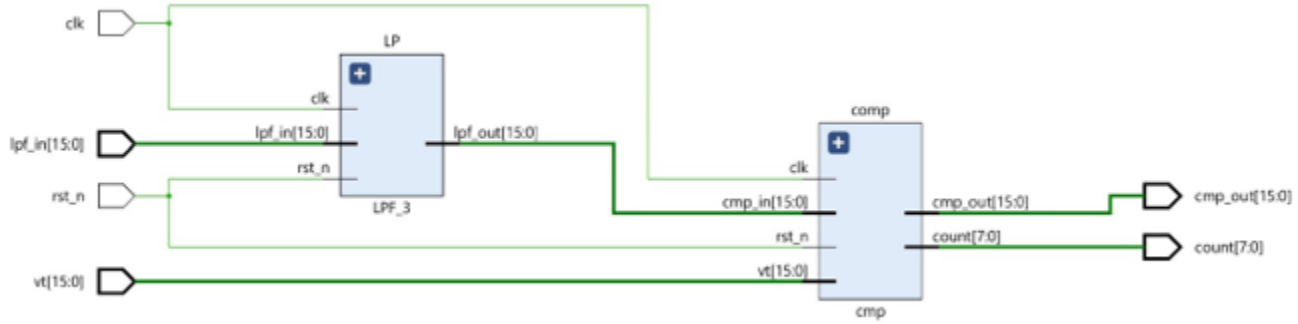
Class of ECG data	Total beats	D_{Acc} (%)	*Se* (%)	D_{Err} (%)	Average *TC*
Normal ECG	75016	99.93	99.95	0.001	1.15 Sec./ record
Ventricular premature contraction	7130	99.88	99.91	0.002	1.25 Sec./ record
Atrial premature beat	2546	99.89	99.82	0.003	1.18 Sec./ record
Left bundle branch block	8075	99.86	99.86	0.002	1.10 Sec./ record
Right bundle branch block	7259	99.92	99.92	0.001	1.07 Sec./ record
Paced beat	7028	99.93	99.93	0.006	1.28 Sec./ record
Fusion of paced and normal beat	982	99.90	99.92	0.004	1.05 Sec./ record

TABLE V
PERFORMANCE OF BETWEEN THE EXISTING AND THE PROPOSED ECG DETECTION METHOD

Performance Metrics	References				
	[35]	[39]	[40]	[41]	Proposed
Average D_{Acc} (%)	99.38	99.86	99.77	99.70	**99.90**
Average *Se* (%)	99.89	99.80	99.60	99.31	**99.90**
Average D_{Err} (%)	0.736	0.34	0.006	0.009	**0.002**
Average *TC* (Sec./record)	1.35	1.28	1.21	1.19	**1.15**



(a)



(b)

Fig. 6. RTL top view of the proposed ECG detector. (a) Schematic of wavelet filter bank used in the proposed detection scheme, (b) Complete schematic of the proposed ECG processing scheme

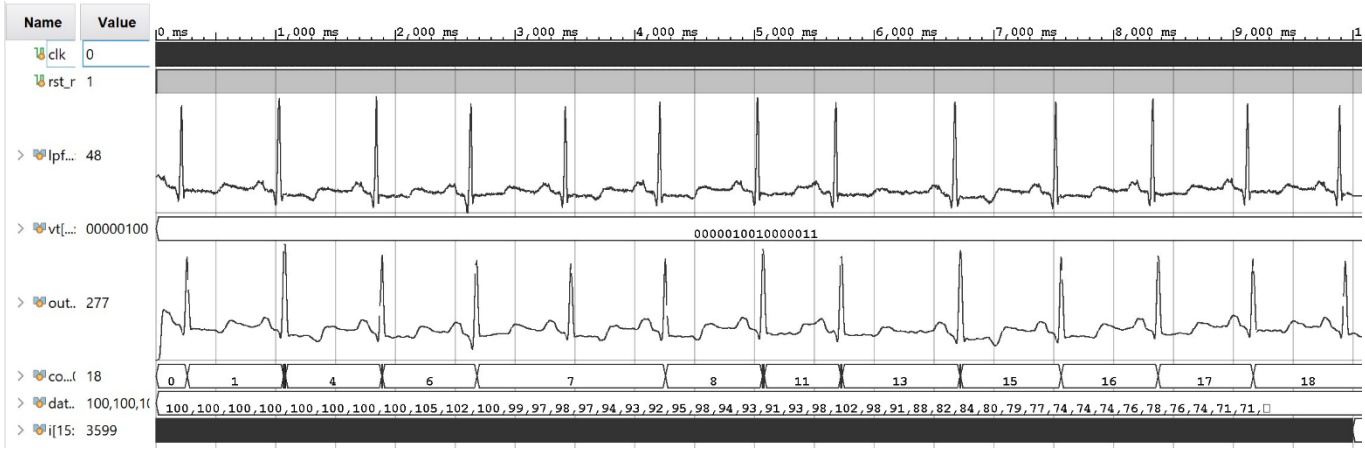


Fig. 7. The output of the proposed ECG detector

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

TABLE VI
AREA, POWER, AND DELAY COMPARISON OF THE PROPOSED ECG DETECTOR WITH EXISTING LITERATURE

Reference	[34]	[35]	[40]	[41]	Proposed
Operating frequency (Hz)	10^6	10^3	10^3	10^3	10^3
VDD (V)	3	3	3	3	3
Power consumption$ (P_S+P_d) (µW)	1420	16.6	0.560	1.9	**0.493**
Area (mm^2)	NR#	1.36	1.19	6.83	**1.1**
Delay (ns)	29.47	31.03	18.73	26.84	**10**
Switching energy (PDP) (µJ)	41847.4	515.098	10.488	50.996	**4.93**

$:$P_S$: Static power, P_d: Dynamic power', # NR: Not Reported

ECG processing scheme is shown in Fig. 6.

Initially, a 16-bit ECG signal, clock, and reset is applied as an input to the wavelet filter bank. The threshold function, which is applied as one of the inputs to the comparator circuit is also of 16-bit. Both online available ECG signals and real-time ECG data recorded using iworx^R IX-TA-220are were used to test the performance of the proposed ECG detector. The output of the proposed detector on a real-time ECG data is shown in Fig. 7. Area, power, and delay are the measures used to compute the computational complexity of the proposed detector. Area requirements of the proposed wavelet filter bank architecture are reduced by using a demand-based wavelet filter bank architecture. Further, LWDF based digital filter structure realization reduces the multiplier's and delay element's count by 75% and 80%, respectively. A fixed-point algorithm, instead of a floating-point algorithm, is utilized to reduce the power requirements of the proposed design. Area, power, delay, adder/multiplier count of the proposed design is shown in Table VI. Table VI brings us to the conclusion that the proposed ECG detector has a lower area, power, delay, and switching energy when compared to the existing work. A fair comparison between the proposed ECG detector and existing work in the literature is made by implementing the proposed algorithm in a similar environment as in work found in the literature to calculate the area, power, and delay requirements.

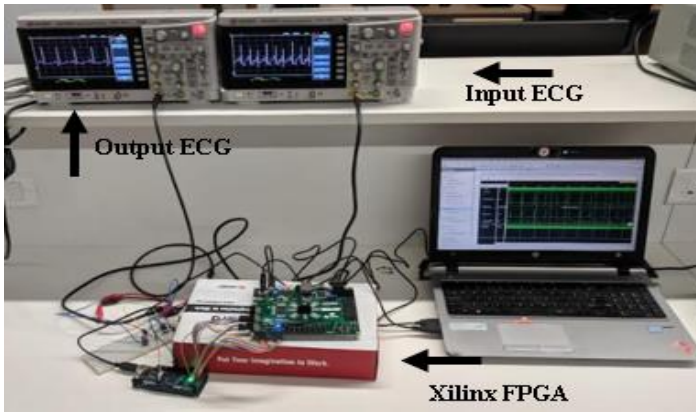


Fig. 8: FPGA implementation and test platform of proposed ECG detector

Fig. 8 shows the FPGA implementation and test platform of the proposed ECG detector.

IV. CONCLUSION

Using biorthogonal 3.1 wavelet transform and an adaptive slope prediction criterion, high detection accuracy, and lowest error has been achieved. Also, using the demand-based wavelet filter bank architecture and lattice wave digital filter realization, the lowest power consumption, area, delay, and switching energy have been achieved. Finally, the real-time usage of the proposed design is validated by hardware implementation on Xilinx-Virtex-7 FPGA.

REFERENCES

[1] Adapted from: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1 (Accessed on 20, August- 2020).

[2] Y.T. Zhang *et al.,* "Challenges and opportunities in cardiovascular health informatics," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 3, pp.633-642, Feb. 2013.

[3] M. Elgendi *et al.,* "Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems," *PloS one*, vol. 9, no. 1, pp.e84018, Jan. 2014.

[4] U. Satija, B. Ramkumar, and M.S. Manikandan, "A review of signal processing techniques for electrocardiogram signal quality assessment," *IEEE reviews in biomedical engineering*, vol. 11, pp.36-52, Feb. 2018.

[5] A. Kumar, R. Komaragiri, and M. Kumar, "From pacemaker to wearable: techniques for ECG detection systems," *Journal of Medical Systems*, vol. 42, no. 2, p.34, Feb. 2018.

[6] S.K. Berkaya *et al.,* "A survey on ECG analysis," *Biomedical Signal Processing and Control*, vol. 43, pp. 216-235, May 2018.

[7] S. Gutta, and Q. Cheng, "Joint feature extraction and classifier design for ECG-based biometric recognition," *IEEE journal of biomedical and health informatics*, vol. 20, no. 2, pp.460-468 Feb. 2015.

[8] A. Burguera, "Fast QRS detection and ECG compression based on signal structural analysis," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp.123-131, Jan 2018.

[9] K. Friganovic *et al.,* "Optimizing the detection of characteristic waves in ECG based on processing methods combinations," *IEEE access*, vol. 6, pp.50609-50626, Sep. 2018.

[10] S. Kiranyaz, T. Ince, and M., Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp.664-675, Aug. 2015.

[11] X. Zhai, and C. Tin, "Automated ECG classification using dual heartbeat coupling based on convolutional neural network," *IEEE Access*, vol. 6, pp.27465-27472, May 2018.

[12] W. Liu *et al.,* "Real-time multilead convolutional neural network for myocardial infarction detection," *IEEE Journal of biomedical and health informatics*, vol. 22, 5, pp.1434-1444, Nov. 2017.

[13] A. Kumar, M. Kumar, and R. Komaragiri, "Design of a biorthogonal wavelet transform based R-peak detection and data compression scheme for implantable cardiac pacemaker systems," *Journal of medical systems*, vol. 42, no. 6, pp.102, Jun. 2018.

[14] A. Kumar, M. Kumar, and R. Komaragiri, "Design of wavelet transform based electrocardiogram monitoring system," *ISA transactions*, vol. 80, pp.381-398, Sep. 2018.

[15] M.A. Kumar, and K.M. Chari, "Efficient FPGA-based VLSI architecture for detecting R-peaks in electrocardiogram signal by combining Shannon energy with Hilbert transform," *IET Signal Processing*, vol. 12, no. 6, pp.748-755, Mar. 2018.

[16] M. Cesari *et al.,* "A new wavelet-based ECG delineator for the evaluation of the ventricular innervation," *IEEE Journal of translational engineering in health and medicine*, vol. 5, pp.1-15, Jul. 2017.

[17] D. Labate *et al.,* "Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ECG: A comparison," *IEEE Sensors Journal*, vol. 13, no.

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3034742, IEEE Sensors Journal

7, pp.2666-2674, Apr. 2013.

[18] Y. Zou *et al.,* "An energy-efficient design for ECG recording and R-peak detection based on wavelet transform," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 2, pp.119-123, Nov. 2014.

[19] J. Pan, and W.J. Tompkins, "A real-time QRS detection algorithm," *IEEE transactions on biomedical engineering*, vol. 3, pp.230-236, Mar. 1985.

[20] A. Singhal *et al.,* "An efficient removal of powerline interference and baseline wander from ECG signals by employing Fourier decomposition technique," *Biomedical Signal Processing and Control*, vol. 57, p.101741, Mar. 2020.

[21] J. Wang *et al.,* "Fractional zero-phase filtering based on the Riemann–Liouville integral," *Signal Processing*, vol. 98, pp. 150-157, May 2014.

[22] N. Razzaq *et al.,* "An intelligent adaptive filter for elimination of power line interference from high resolution electrocardiogram," *IEEE Access*, vol. 4, pp.1676-1688, Mar. 2016.

[23] M.N. Salman, P.T. Rao, and M.Z.U. Rahman, "Novel logarithmic reference free adaptive signal enhancers for ECG analysis of wireless cardiac care monitoring systems," *IEEE Access*, vol. 6, pp.46382-46395, Aug. 2018.

[24] D.P. Tobón, and T.H. Falk, "Adaptive spectro-temporal filtering for electrocardiogram signal enhancement," *IEEE Journal of biomedical and health informatics*, vol. 22, no. 2, pp.421-428, Dec. 2016.

[25] M.X. Cohen, *Analyzing neural time series data: theory and practice*. MIT press, 2014.

[26] A. Kumar, R. Komaragiri, and M. Kumar, "Heart rate monitoring and therapeutic devices: a wavelet transform based approach for the modeling and classification of congestive heart failure," *ISA transactions*, vol. 79, pp.239-250, Aug. 2018.

[27] A. Sharma *et al.,* "Accurate tunable-Q wavelet transform based method for QRS complex detection," *Computers & Electrical Engineering*, vol. 75, pp.101-111, May 2019.

[28] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 1, pp.21-28, Jan. 1995.

[29] S. Kadambe, R. Murray, and G.F. Boudreaux-Bartels, "Wavelet transform-based QRS complex detector," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 7, pp.838-848, Jul. 1999.

[30] Adopted from: http://www.iworx.com/documents/brochures/IX-TA-220.pdf (Accessed on 20, August- 2020.)

[31] Adopted from: https://physionet.org/ (Accessed on 20, August-2020.)

[32] M. Abdolmaleki *et al.,* "Selecting optimum base wavelet for extracting spectral alteration features associated with porphyry copper mineralization using hyperspectral images," *International Journal of applied earth observation and geoinformation*, vol. 58, pp.134-144, Jun. 2017.

[33] A.L. Goldberger *et al.,* "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp.e215-e220, Jun. 2000.

[34] S. Talukder *et al.,* "An Efficient Architecture for QRS Detection in FPGA Using Integer Haar Wavelet Transform," *Circuits, Systems, and Signal Processing*, pp.1-16, Jan. 2020.

[35] Y.J. Min *et al.,* "Design of wavelet-based ECG detector for implantable cardiac pacemakers," *IEEE transactions on biomedical circuits and systems*, vol. 7, no. 4, pp.426-436, Feb. 2013.

[36] J.S. Sahambi, S.N. Tandon, and R.K.P. Bhatt, "Using wavelet transforms for ECG characterization. An online digital signal processing system," *IEEE Engineering in Medicine and Biology Magazine*, vol. 16, no. 1, pp.77-83, Jan. 1997.

[37] H. He, Y. Tan, and Y. Wang, "Optimal base wavelet selection for ECG noise reduction using a comprehensive entropy criterion," *entropy*, vol. 17, no. 9, pp.6093-6109, Sep. 2015.

[38] E.Y. Hamid, R. Mardiana, and Z.I. Kawasaki, "Wavelet-based compression of power disturbances using the minimum description length criterion," In *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No. 01CH37262)*, 2001, pp. 1772-1777.

[39] J.P. Martínez *et al.,* "A wavelet-based ECG delineator: evaluation on standard databases," *IEEE Transactions on biomedical engineering*, vol. 51, no. 4, pp.570-581, Mar. 2004.

[40] P. Li *et al.,* "A 410-nW efficient QRS processor for mobile ECG monitoring in 0.18-µm CMOS," *IEEE Transactions on biomedical circuits and systems*, vol. 11, no. 6, pp.1356-1365, Aug. 2017.

[41] C.I. Leong *et al.,* "A 0.83 µW QRS detection processor using quadratic spline wavelet transform for wireless ECG acquisition in 0.35 µm CMOS," *IEEE Trans. Biomed. Circuits Syst*, vol. 6, no. 6, pp.586-595, Dec. 2012.

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:36:17 UTC from IEEE Xplore. Restrictions apply.

Volume 8 Issue 2 (2020) 86-98 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Sunil, Dharmanshu Kumar, Udit Mehta, Rakesh Kumar
Department of Mechanical Engineering, Delhi Technological University, Bawana Road, New Delhi-110042, India

Article Info

Article history:
Received 17 January 2020
Received in revised form
20 May 2020
Accepted 28 May 2020
Available online 15 June 2020

Keywords: RSC, Barrier, Fuzzy, Critical success factors

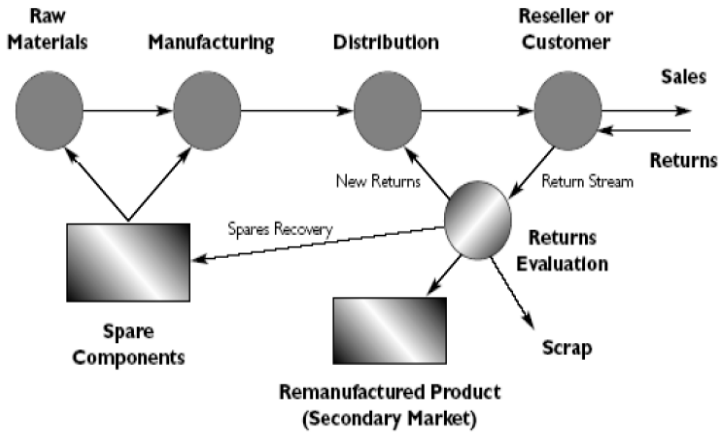
Abstract: Reverse supply chains represent an opportunity to create a value stream, not financial loss. Reverse supply chains deserve as much attention at the corporate level as forward supply chains and should be managed as business processes that can create value for the company. Many manufacturing firms have not yet adapted reverse supply chain but view it as the great nightmare to their organizational system. Like many other system, reverse supply chain has also its critical success factors. This research aims to identify the barriers of reverse supply chain management which invariably present in the system. Among the critical success factors, those perform below the threshold were identified as barrier using fuzzy logic technique.

1. Introduction-Reverse supply chain management

Retrieving the product from customers end this means end of the forward supply chain processes. It is just the safe and secure ways of challenges for the reverse supply chain that need implementation on a easy way so that people can easily trace the technology those who facing difficulties and and from unknown barriers, this paper is aims to works on its critical success factors barriers of RSCM and to go on a next level of Implementation on customer and corporate level to make it more barrier free environment in upcoming researches as reliable technology in this era.[1] To respond the increasing demands and decreasing life cycle of products, what companies are doing is appliying more RSCM in their firms to optimize their business effectively and increase their overall performance profitability in their firms. The streamlined task tool which optimally collecting end-of-life (EoL) products, repair them and sell them again to their customer this entire things referred as Reverse Supply Chain Management (RSCM)[12]. Reverse supply chain management is used to integrate effieciently, from various manufacturers, warehouses, retailers, and stores, so can help in merchandise and distributed in a right way, in order to minimize costs while satisfying service level requirement at both forward and reverse same chain. RSCM increases importance as a profitable and sustainable business strategy.[4] This chain describes a holistic view of reverse supply chain and presents its main attributes, success factors, barriers and other useful professional tools.A company faces problem when it has to manage the reverse supply chain efficiently. Unlike Supply chains, such important critical factors are yet not discoverd and developed for the reverse supply chain. As both supply chains have their different features and importance and cannot be treated as comprarable to each other[5]. In India, as Underdeveloped countries various things are unexplored the experience of the reverse supply chain is relatively new and there is lack of knowledgeand adaptation of this chain but in developed coutries they are started to implement this or this process is forwarded as in their limitations. Limited number of factors affecting reverse supply chain and availibility of these factors can become barriers of reverse supply chain implementation in an industry due to various rules and regulation or behavior of peoples in various countries.. Enumerated strategic factors government rules and regulations, consumer concerns, environmental issues and costs and with other operational factors such as cost-benefit analysis, logistics, warehousing, supply chain management, recycling, and packaging. Agrawal et al. [5] identified some major critical factors in the Indian industry a management awareness, resource management, economic factors, and contracts terms & conditions. These factors help the firms to take suitable action for adaptation and implementation.

Other cases of products direction in the supply chain are manufacturing returns,(B2B and B2C), product recalls, warranty returns, repair returns, returns and end-of-life products are reversed from the original supply chain. Reverse supply chains are categorised, and they lie in the different stages of the product life

Corresponding Author,
E-mail address: uditmehta12@gmail.com
All rights reserved: http://www.ijari.org



Reverse supply chain management

cycle; however, most chains are organized to carry out five key processes: (cognizant 20-20 insights | January 2011)

Product acquisition: Acquisition of products

- Product acquisition: Acquisition of products
- Reverse logistics: Processing logistics from company and customer.
- Inspection & disposition: Assessing the condition of the return and making the most profitable decision for reuse.
- Reconditioning: Recast into their original specifications
- Distribution & sales: Creating secondary markets for the recovered products

Once the products are analysed into different categories, they are disposed into other reuseable forms like reuse, or repairing, or remanufacturing, or recycling depending on factors [3]. Treatment should be industries or product specific and depends upon the number of factors including price/value / shelf life of the product, transportation cost, and the demand outlook. The consumer expects replacement of product as soon as possible while the manufacturer wants to recapturing the maximum value from it.[5].

Typical examples of reverse supply chain include: Printing industry,Computer industry,Automobiles,Electronic industry. The problem arrives when a company has to minimize its cost efficiently. Unlike forward supply chains, critical factors for the reverse supply chain are undeveloped and unexplored. Both have different characteristics and cannot be treated by the same methodology. Furthermore, there are many factors and problems in the implementation in the reverse supply. , reverse logistics plays a very important role in company's competitive advantage and help strategic decision making[12]

In this we see various barriers in reverse supply chain management and used to identified its critical factors and work on it to enhance the efficiency of any organisation reverse supply chain activities by reknowned fuzzy technology and use to rank critical factors in upcoming topics by clearly defined the topics.

2. Reverse logistics

Reverse Logistics is the kind of activity which led after an offer for an item to recover worth and end lifecycle of the item. It is a new pattern of enterprise strategy management and supply chain management. RL pursue a coordinated optimization of economic and social efficiency of the product. It is not only the foundation of a circular economy but also had become the synergism with the development of a sound domain. It is the way toward moving merchandise from their place of utilization to the point of root to catch esteem or legitimate removal. Turn around coordination is once in a while called secondary selling flexibly chain, reseller's exchange coordination or logistics. The natural and asset results of an item are surveyed and assessed through Life Cycle Analysis (LCA) through all the stages (removing and handling the crude material, creation, dispersion, use, remanufacturing, reusing and last removal) of its life. Ventures may utilize LCA to help item improvement with the goal that the general ecological effect of the item can be limited. The procedure includes moving returned merchandise to the area of recuperation or removal legitimately or thought testing, assessment and arranging offices (Guide and van Wassen hove, 2002; Krikke et al., 2004). The primary converse coordination activities are transportation, warehousing and stock administration.

In Reverse logistics the processes which have to undergo for the product consists:

* Remaking - revamping the item to reuse, fix or to make new parts
*Recondition - The general classification which incorporates client support, the administration and the return of the product
* Return management
* Management for waste and Recycling
* Management for Warranty
* Management for Warehouse

As indicated by Tibben Lembke and Rodgers (2002), turn around coordination is a basic issue for some organizations and records for four to 10% of all out coordination cost. Hence, the advancement of converse coordination framework and the viable administration of opposite coordination exercises give a likelihood to decrease cost, increment incomes, and create extra benefit (Poist, 2000) As indicated by Blackburn et al. (2004) because of various qualities of profits and various prerequisites for various sorts of profits (lead-time, costs), invert gracefully the chain normally comprises of a few converse coordination streams with various properties. Wu and Cheng (2006) expressed that there is an essential distinction in the objectives and necessities among forward and switch coordination. As indicated by Fleischmann et al. (1997) when contrasted with the forward coordination' streams the retrogressive streams typically experience the various streams. Moreover, various kinds and attributes of profits confuse operational parts of converse coordination (Guide et al., 2003).Then again, financial reasons have additionally had their commitment to this expanding significance of Reverse Logistics issues. Using returned items, organizations stand the chance of recuperating either constituent material [7]. For instance, produced products could go from customer to merchant or producer. These products, materials, and gear could be remanufactured, reused, recuperated, or reused. A great many PCs on the planet got obsolete. There are chances to reuse these PCs and make new qualities. E-squander contains PCs, TV sets, sound gear, batteries, and so on. E-squander contains aluminum, copper, plastics, glass, and so on. Inside the car business, waste could incorporate motors, alternators, starters, transmissions, and so on. Turn around coordination is as of now one of the most significant issues identified with PC makers, printers, car industry parts, shopper hardware, family unit synthetic compounds. There are different reasons why return streams exist. They could be identified with client support, ecological or monetary. Enormous world organizations effectively fit for taking care of converse coordination issues incorporate DHL, UPS, FedEx, and so on. Turn around coordination is getting significant in the cutting edge world, because of the expanding regard for the insurance of the earth. Clients and customers are facing difficulties in their logistics handling having barriers in their implementation and solving this issue using various tools [9].

For any organization, key factors have a drawn out main concern sway. These factors must be overseen adequately, effectively and proactively — not strategically or responsively — for the feasibility of the organization. As of not long ago, most organizations focused uniquely on business capacities, for example, account or advertising as the vital factors. Coordination abilities were taken a gander at as key factors during the late 1970s and 1980s. Numerous organizations consider invert coordination as holding a significant key job, however this capacity still can't seem to pick up the status of a vital variable. The significance of opposite coordination is expanding for a few reasons expressed beneath:

a) Companies are seeing substantial advantages from a worth that can be recovered from ineffective resources coming about because of returned stock, for example, noteworthy decreases in inventories, improvement in income, diminished work and improved consumer loyalty.
b) Competitive weight has expanded to give a successful, effective returned products process. The expansion in inventory and e-business shopping has brought about the advancement of merchandise exchanges to pick up client trust and diminish chance.
c) The pressure of the item life cycle and an expanded accentuation on presenting new items have made a need to clear the appropriation channel all the more as often as possible, requiring the productive methods which brings back old things.
d) Increased administrative necessities identified with reusing and air of items—particularly for the items which are risky for the earth—have expanded the requirement for exactness record-keeping and following [3]. The case of Reverse coordination is Automotive Industry-The car business is perhaps the biggest business on the planet and manages the most costly of buyer products. In this way, it isn't astounding that opposite coordination is a significant subject for this industry. Reverse coordination assumes a critical job in three essential regions:
e) Recovering of goods and materials from the finish of life vehicles.
f) Stock adjusting returns of the new parts from sellers. The huge three automakers in the U.S. have combined to frame the Vehicle Recycling Development Centre to expand the conscious 20-20 bits of knowledge 3 aware 20-20 bits of knowledge 4 recyclability of vehicles. At VRDC, the emphasis is on figuring out how to assemble vehicles that can be dismantled all the more without any problem. The middle is researching one of the freshest patterns in building which is Design for Disassembly (DFD). From this step the dismantling of the product become simpler by just reducing the number of parts, fittings. In contrast to other ecological activities for assembling, DFD offers the chance of numerous unintended constructive outcomes, for example, remanufacturing. The car business might be the business with the longest history of utilizing old items. As per the Auto Parts Remanufacturers Association, the market for automobile parts that are remanufactured is evaluated at around $34 billion for every year. Additionally gauges that there are 12,000 remanufacturing firms which are engaged with the vehicle parts industry. A specific organization remanufactures in excess of 4,000,000 alternators, starters, and water siphons every year. Between, 90% to 95% of the considerable number of starters and alternators which are sold for substitution are remanufactured. Automakers need to keep up a shut circle framework by their parts. In the event that a vehicle needs another transmission, it is trusted that the shopper will carry the vehicle to the vendor that will supplant the old transmission with a remanufactured one. The seller will send old transmission to the automaker for remanufacturing. Along these lines, the

automaker will keep up a steady gracefully of transmission centers. The equivalent SCM and internet business advances engaged with moving items to buyers (known as forward coordination) are utilized backward coordination, including standardized identifications and scanners used to follow returns, materials taking care of frameworks in distribution centers and the exchange of electronic data records between gracefully chain suppliers. SCM and ERP programming merchants were at first delayed to help switch coordination, as indicated by certain specialists, however most dealers currently incorporate some opposite coordination includes in their suites.

3. Importance of Revesres supply chain management

The importance of reverse supply chain is increased in past few years by the peoples:

- Facing high return in some industries having more than 50%.
- increased sales opportunity
- End-of-life take-back regulations have flourished over the past decade demanding businesses to effectively bring about the entire life of the product.
- Consumers and government successfully pressured businesses to do something or take responsibility for disposals of their EOL products that are harmful to earth and living also.
- Landfill limited capacity and highly luxurious. Some items like, remanufacturing and recycling has become more predominant and feasible.

Fig.2: Reverse Supply Chain

- Return on Investment
- Increase in Public Perception
- Reuse Encourages Competition in Manufacturing

3. Literature review

3.1. Reverse supply chain management

The reverse supply chain is the forwarding goods from the end customer to the manufacturer again. Reverse logistics refers planning, implementing and controlling the efficient and effective flow of secondary goods and information.

Instances of the converse gracefully chain include:

- Remanufacturing and renovating
- Item return and the board of testimony
- The arrival flexibly chain is having five key procedures as:
- Item securing: This is the way toward acquiring the pre-owned item from the client by the producer.
- Invert coordination: This is the shipping items to an office for reviewing, arranging and mien.
- Investigation and mien: This incorporate the way toward surveying the returned item and settling on the choice for reuse.
- Remanufacturing/Repairing: It is the way toward making the returned item to its unique determination and capacities.
- Advertising: Making the market for recouped and remanufactured items.

Figure 3, shows invert flexibly chain for business item returns. The client restores the pre-owned item to affiliate (item obtaining), from where the items are dispatched to assessment focuses (turn around coordination's) and afterward assessment needs to happen (examination and attitude). Evaluator's tests are to be performed to decide the choices of brought items back.

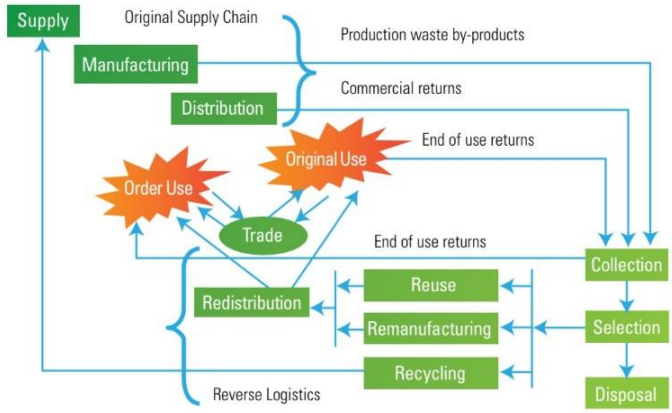


Fig.3: Reverse supply chain management

3.2 Centralized Efficient Reverse Supply Chain

A proficient converse flexibly chain centres around the handling of the arrival of items at a general lower cost. This flexibly chain model penances speed over cost proficiency and is commonly material to items that have shorter time/esteem devaluation. The retailer or affiliate doesn't partake in any item assessment at their end. The transportation cost minimization is furnished by the profits related to the delivery of the product in mass.

When the state of the item is resolved at the testing office, it is suitably discarded (i.e., sent for restocking, revamping, rescuing or rejecting). This model is utilized to accomplish economies of scale at each level, consequently empowering minimization of handling costs. It is likewise simpler on the retailer, which doesn't need to sort returns or boat them to different areas, just as the outsider credit guarantor, who needs to give the credit at once for various item returns.

3.3 Decentralized Reverse Supply Chain with Preponement

In the past model, it was seen that item separation is deferred until all profits are transported to the focal office. This can be thought of as a deferment system. On the other hand, for the items like PCs where the benefit esteem misfortune after some time is high, we can prepone the testing and assessment of the returned item utilizing a decentralized model with the goal that the unused items can be restocked right away. This testing can be performed at the purpose of return by the affiliate or the retailer, however it would should be in fact achievable, and it may require mastery advancement among affiliates. Besides, motivator would be given to affiliate through shared reserve funds contracts with the makers, or merchants could set up a seller oversaw stock procedure with enormous retailers and keep up their own specialists to test the brought items back.

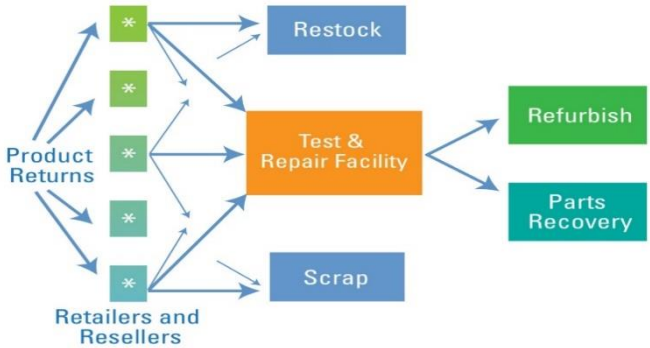


Fig.4: Reverse Supply Chain with Preponement

The alternatives are identified here is to improve these all alternatives to improve in industrial applications and further research Environment-friendly, reliable and Consistent product acquisition, inventory reduction, product analysis, reverse distribution, cost minimization, and criteria are analysis from various peer research

papers to obtain barriers in implementing successfully reverse supply chain management and identifying all critical factors in reverse supply chain which impact more and give them ranking based on fuzzy logic, here survey is analyzed, to appropriate use of coordination mechanism which expected to increase efficiency and effectiveness in the operation, the actors and coordination members, therefore the selection of criteria impact the performance of reverse supply chain management. Because of the multi-dimensional criteria, the selection of appropriate RSCM in a given situation remains a difficult task for managers. This paper is an endeavor to investigate different issues of opposite flexibly chain coordination and utilizations an incorporated methodology of Fluffy AHP and Fluffy TOPSIS to organize coordination components and afterward organize exchanging accomplices which depend on coordination instruments. [61]

To oversee unclearness and vulnerability in dynamic, Zadeh (1965) proposed a Fluffy set hypothesis. When displaying is finished utilizing Fluffy sets it has end up being an effective path for definition of choice issues, where the data accessible is emotional and loose (Zimmermann, 1992). The Fluffy numbers represent specific go for specific esteem. Because of this specific run, it is simpler for the evaluator to demonstrate his/her inclination. The inclination of the master is in numerous down to earth cases is unsure, which makes it difficult to make a numerical examination (Torfi, Farahani, and Rezapour, 2010). So, a solitary etymological rating will be converted into a Fluffy number comprising of various numbers. Along these lines, the phonetic rating is reflected as a range. Both triangular and trapezoidal fluffy numbers can be utilized for the Fluffy hypothesis however we are restricted here to triangular fluffy numbers. Balli and Korukoglu (2009) keep up that it is frequently helpful to utilize triangular Fluffy numbers (TFNs) as a result of the simplicity of computation.[61]

$$\mu_A(x) = \begin{cases} 0 & x < l; \\ \frac{x-l}{m-l} & l \le x \le m; \\ \frac{m-x}{u-m} & m \le x \le u; \\ 0 & x > u. \end{cases}$$

Deng (1999) discusses this mathematical representation of a TFN M that is depicted by Balli and Korukoglu (2009) as shown in Figure 1. Next sections present an overview of both techniques including important steps and previous applications of these in the research work

3.4 Identifying goals and barriers of Reverse supply chain

After reviewing paper of peers, we have identified barriers in reverse supply chain which can cause certain issues while implementing this in the system we also have identified our alternatives as:

i. Environment-friendly
ii. Reliable and consistent product acquisition
iii. Inventory reduction
iv. Product analysis
v. Reverse distribution
vi. Cost minimization

These have been identified as the goals in reverse supply chain, so we need to find the priority order in this to see the most prominent way to find the critical factors in RSC. These are the major objectives that enhance the performance of RSC in implementing in any industry. The description of these factors is written as-

3.4.1 Environment-friendly

It has the topmost order from the objectives in today's scenario as we need to focus on our environment because it comes first in every aspect of technology. Reverse supply chain is the part of Supply chain which deals with the product return. Waste management, recycle and reuse, etc. related to the environment. This is the reason why researchers are preferring this potential area to observe how the reverse supply chain is can be used to enhance and protect the environment.

3.4.2 Reliable and consistent product acquisition

Product acquisition is exceptionally important in as the behaviour of the consumer and is very sensitive, how the product is recovering in stages whether it comes as repair or unused stage. In this paper, we want to shed light on the process consumers go through when making acquisitions by exploring the company acquisition process.

3.4.3 Inventory reduction

Stock spares our business conveying costs, stockpiling expenses, and transportation costs between stockroom offices. Stock decrease takes out of date stock, which if not sold/recoup under critical conditions, will go to finish waste and income down the channel.

3.4.4 Product analysis

The analysis is a very important factor in RSC. After all, it creates a benefit in the organization because it defines which product is going for disposal, reuse, recycles, secondary market because it recovers some revenue for this process and makes benefit to the company.

3.4.5 Reverse distribution

Conveyance has effect the general public it restores the item in the hands of a client or optional market it additionally the RSC and make a tremendous factor in recognizing the hindrances of RSC, it for the most part makes the procedure quicker and trustable according to clients.

3.4.6 Cost minimization

Minimization of cost is a strategy that is to achieve the most cost-effective path for delivering the goods and services to appropriate quality. It is an important aspect to consider that cost minimization is not about the degradation of quality; rather it is about to meet customers' needs to make a profit in the process. So, these some objectives/alternatives in RSC to which the attributes are taken.

The identification of these barriers is taken from various peer's review literature and other sources and from experts who are studying the reverse supply chain most prominently from the top organization of India.

3.5 Barriers

3.5.1 Environment driven legislation

Legislation refers to regulations or acts passed by the government authorities to minimize the effect of the end of life products on the environment. Cause in every state we are offering different rules and regulation of product, focusing on environmental concerns is partly enforced by government legislation (Prendergast and Pitt 1996). The government of India has established e-waste Rules (Management and Handling), which have come into effect from the year 2012. Experts mentioned that sooner or later Indian manufacturers will have to comply with these waste management regulations.

3.5.2 Provide facilities for assembling/disassembling

We don't know the nature of products so the company must have these types it becomes a barrier as it increases the cost of this process, so the company looks for an alternative solution for this process.

3.5.3 Revenue management

Revenue management is the application of disciplined analytics that predict product at the micro-market levels and optimize product price recovery to maximize revenue growth. The main goal of revenue management is selling the right product to the right customer at the right time with the right packaging. In this field of work, the company has to get the right perception of the customer about the product and deciding the value price accordingly and make sure about the appropriate availability of the product.

3.5.4 Customer awareness and responsibility

For the return of EOL products the process in the customer is very important cause most of the customers are don't want to return the product and after so long they make the products orphan.in that case, the company not recognized its analytics process of recovery, so, customer awareness is a very critical factor in this procedure.

3.5.5 Waste Management

Waste management involves taking difficult decisions at strategic and operating levels.. Disposal may include transportation, landfilling, and incineration steps. Disposal is required for products

that cannot be reused for technical or economic reasons or they are completely consumed.

3.5.6 Forecasting return

In the wake of investigating obtainment choices, scope organization, and removal the executives. At an each operational level, nitty gritty forecasts of the amounts to be returned in every period, just as the changeability of these amounts, is valuable, particularly for stock administration and creation arranging.

3.5.7 Location routing

It incorporates the issue of coordinating the area choices of dispersion focuses (DCs), the approaches of opened appropriation places, and the vehicle directing choice in serving clients, in which new merchandise are to be produced and harmed products are to be fixed by the assembling unit and afterward must be shipped to open dissemination habitats (DC).Vehicles that start from and end in the same DC distribute recovered goods to satisfy the demands of customers and retrieve damaged goods. The ultimate goal is to make the total cost of manufacturing and remanufacturing the goods to a minuscule level, establishing new DCs, recovery of the goods between manufacturer and DCs and appropriate distribution.

3.5.8 Network structure of EOL

The stressed stock control of returned streams and converse appropriation, yet in addition featured the absence of a general structure and numerical model to help the opposite coordinations condition. Guide et al. (2000) checked on the writing on stock frameworks with returns and Guide (2000) recognized and portrayed seven confounding qualities of creation arranging and control exercises for remanufacturing firms..

3.5.9 Integration of reverse logistics activities in reverse supply chain

To minimize the cost they proper need of integration of RL into RSC but due to this the problem is facing companies to double spending barriers.

3.5.10 Coordination of information

Proper coordination is required for the successful recovery of EOL products.

3.5.11 Inventory management

The objective of inventory management is to control external component orders and the internal component recovery process to guarantee a required service level and to minimize fixed and variable costs.

3.5.12 Coordination between company and customer

For successful implementation of RSC proper coordination between C2C is needed to overcome barriers.

4. Research Methodology

4.1 Analytics hierarchy process (AHP)

AHP is a multi-rules dynamic technique which is valuable to structure the choice issues dependent on pairwise examinations in the wake of getting phonetic based information from master's decisions in fuzzyfied structure. The AHP found by Saaty in 1980 out of one of the most well-known multi-models dynamic strategies. In association numerous issues couldn't be comprehended on account of fragmented or no accessibility of data. This strategy consolidates the pairwise examination lattice of chiefs' choices and makes a Fuzzy set to deal with the vulnerability. This Fuzzy AHP utilizes the pair-wise examination framework with the assortment of (n-1)/2 correlations. As the quantity of quality and choice inquiry builds, the pair-wise correlation questions and the multifaceted nature of the grid increment and plausibility of respondents answering with incorrect data, which makes irregularities in the outcomes despite the fact that the consistency proportion may not be under 0.1 textures must be check before heading off to the arrangement.

An appropriate connection can be utilized to conquer this issue and for the adequacy in choices and a pragmatic dynamic procedure can be made.

The inclination relations in the AHP improve the consistency of Fuzzy AHP. With the assistance of CFPR, the lessening the quantity of pairwise correlations from n(n-1) to (n-1) examinations figured through the Fuzzy information. This makes procedure increasingly proficient, and leaders require less exertion to consideration more on making the pair-wise correlations of properties. Subsequent to deciding consistency we move to organize the components under same procedure follows in TOPSIS which used to rank the elective significance from the Euclidean separation at this stage it organize the variables and give an appropriate motivation to the business to actualize.

4.2 Fuzzy TOPSIS technique

TOPSIS stands for The Technique for Order Preference by Similarity to Ideal Solution was proposed first by Hwang and Yoon (1981) and Fuzzy TOPSIS strategy by Chen and Hwang (1992). The essential thought for Fuzzy TOPSIS is to pick the other option, as indicated by the closeness level of FPIS and FNIS beyond what many would consider possible. The positive perfect arrangement is an augmented arrangement.

Advantage elective and limited cost rules. In FNIS the cost standards are boosted and advantage models are limited. As indicated by Chan and Kumar (2007), there are some genuine circumstances where human inclination is questionable and leaders may be hesitant or incapable to allot fresh qualities to their decisions. The Fuzzy TOPSIS is capable in managing multi-measures dynamic by appropriate semantic qualities into Fuzzy numbers and in this way permitting leader to consolidate fragmented or inaccessible data into the choice model. Fuzzy TOPSIS to rank the elective request to the client's inclination. In true information are not all that deterministic, it is loose or Fuzzy in nature. Along these lines, an all-encompassing TOPSIS for Fuzzy information as proposed by Chen (2000) and Chen et al. (2006) is talked about underneath.

The calculation of the Fuzzy TOPSIS technique can be portrayed as follows:

Stage 1: Form a group of master to assess the other options.

Stage 2: Identify the assessment models.

Stage 3: Choose the fitting semantic factors for assessing options regarding recognized rules.

Stage 4: Determine the collected load of choices as for every standard. The Fuzzy rating are portrayed as TFNs:

Where,

Rk = (ak, bk, ck) k =1; 2; 3… . K, Fuzzy rating decided as R = (a, b, c), k = 1, 2, 3,... K. Here,

$$a = \min_k(a_k), b = \frac{1}{K}\sum_{k=1}^{K} b_k, c = \max_k(c_k).$$

Stage 5: make the Fuzzy choice framework

Stage 6: Normalized the Fuzzy choice framework. For standardization, the straight scale trans-development can be utilized to change the different rules scales into an equivalent scale. We can get standardized Fuzzy choice network R~ (Chen, 2000).

$$\tilde{R} = [r_{ij}]_{m\times n} \quad i = 1, 2, 3, \ldots, m; \quad j = 1, 2, 3, \ldots, n$$

$$\text{where } \tilde{r}_{ij} = \left(\frac{a_{ij}^*}{c_j}, \frac{b_{ij}^*}{c_j}, \frac{c_{ij}}{c_j^*}\right) \quad \text{and } C_j^* = \max_i C_{ij}.$$

Stage 7: Construct the weighted Fuzzy matrix which is standardized.

$$V = [\tilde{v}_{ij}]_{m\times n} \quad i = 1, 2, \ldots, m \quad \text{and } j = 1, 2, \ldots, n$$

$\tilde{v}_{ij} = \tilde{r}_{ij}W$, where W is the weighted vector of evaluating criteria.

Stage 8: Determine positive ideal and negative ideal solution

$$\text{FPIS(P*)} = (\tilde{V}_1^*, \tilde{V}_2^*, \tilde{V}_3^*, \ldots, \tilde{V}_n^*) \quad \text{and FNIS}(P^-) = (\tilde{V}_1^-, \tilde{V}_2^-, \tilde{V}_3^-, \ldots, \tilde{V}_n^-)$$
$$\text{where } \tilde{V}_j^* = \max_i\{v_{ijk}\} \text{ and } \tilde{V}_j^- = \min_i\{v_{ijk}\}; i = 1, 2, \ldots m; j = 1, 2, \ldots, n.$$

Stage 9: Find the distance between positive and negative ideal solution

$$d_i^* = \sum_{j=1}^{n} d_v(\tilde{v}_{ij} v_j^*); \quad i = 1, 2, 3, \ldots, m \quad \text{and } d_i^- = \sum_{j=1}^{n} d_v(\tilde{v}_{ij} \cdot v_j^-); \quad i = 1, 2, 3, \ldots, m$$

i= 1,2,3…. where, (dv) = distance of both

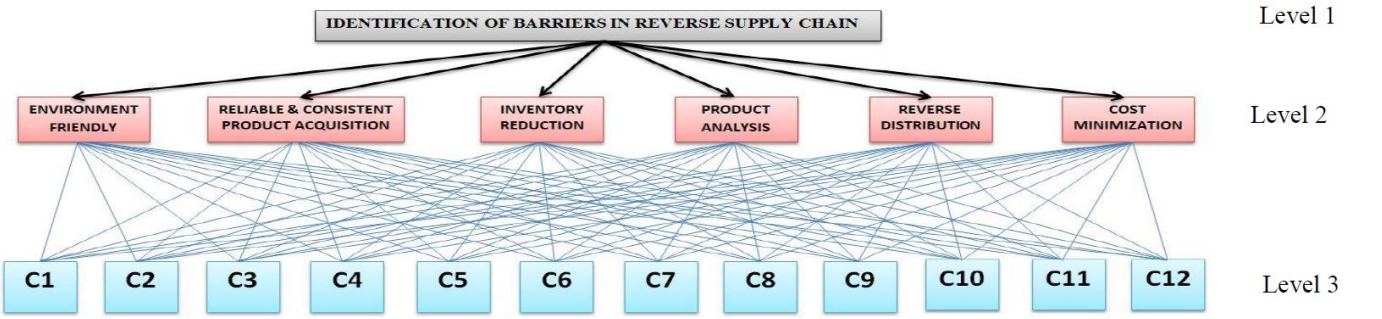


Fig.5: Barriers in Reverse Supply Chain

4.3 Fuzzy numbers

Step 10: Calculating the closeness coefficient for each alternative Closeness Coefficient (CCi) is defined as rank alternatives. The closeness coefficient represents the distance to the FPIS (P*) and Fuzzy negative ideal solution (P−). By using this formula we measure closeness coefficient for each alternative:

$$CC_i = \frac{d_i^-}{d_i^- + d_i^*}, \quad i = 1, 2, 3, \ldots, m.$$

Step 11: Ranking according to the closeness coefficient, the ranking of the alternatives can be determined. According to equation, would be closer to FPIS and farther from FNIS as CCi approaches to one.

4.4 The Proposed model

The structure of the model is shown in Fig.1 taken from our traits and afterward choices, at first arranged for the dynamic procedure by means of AHP. Be that as it may, before the database is gathered from peer looked into writing papers, specialists' decisions and the indexes of various producers. Some gathering are taken for getting a few thoughts regarding venture from specialists for the other options and properties.

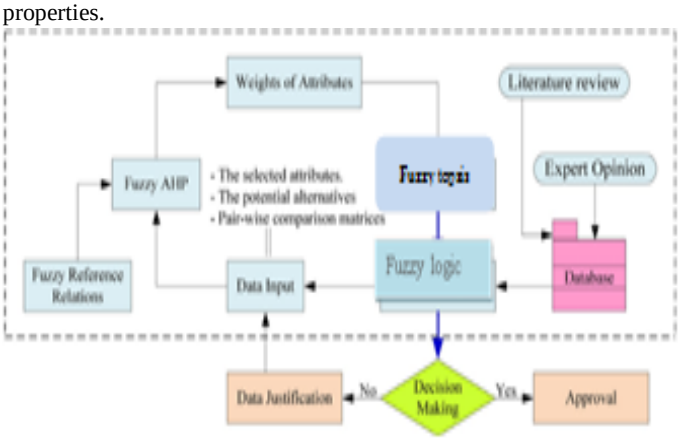


Fig 6: Scheme of the proposed model.

The understanding further information inputs we utilized the fluffy AHP with the inclination relations. The rules loads of characteristics are determined by the improved fluffy AHP with the pair-wise examination lattice dependent on the specialists' decisions and fluffy inclination relations. The consequences of the improved fluffy AHP are the inputs of the fuzzy topsis for determining the ranking of alternatives. After analyzing the data results, rank the variables. In the event that the outcome is agreeing from AHP or on the off chance that not, at that point information support ought to be completed for contributions to improved fluffy AHP and after conclusive outcomes might be reported.

The qualities in the choice help model are extricated from the writing, inventories and from converses with specialists in assembling. It contains three top-down levels: At the first level (level 1), the reverse supply chain goal is determined for critical factors: the middle level (level 2) consists of attributes for the decision- making process, level 3 alternatives such as The RSCM alternatives are as (A1, A2, A3, A4, A5, A6) are listed in the bottom level (level 3) for the ranking process.

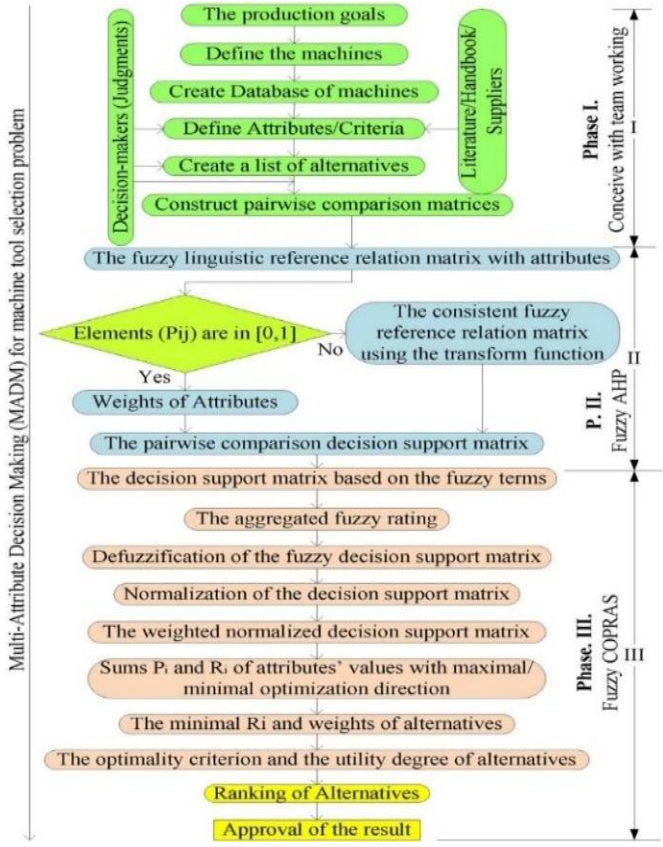


Fig 7: Activities series

4.5 Priority weight for decision criteria

After findings of all barriers in the literature we all need to do prioritize all factors on basis of alternatives or attributes of using Fuzzy AHP is to determine important weight of the criteria that will later in Fuzzy TOPSIS method. As per the process we do pairwise comparison by matrix that matches linguistic statement of data is formed by our decision makers to fill it. A group matrix will be obtained by calculating average of Fuzzy numbers for all samples. Following steps explain the method of determining priority weights for decision criteria: which are described [14]

Step 1: A panel of three decision makers from the organization is selected as per their experience in the area of reverse supply chain management and role in the organization.

Step2: Environmental-driven legislation, provide-facilities-disassembling, revenue-management, customer-awareness and responsibility, waste management ,forecasting return ,location routing, network structure of eol, integration of rl activities into rsc, coordination of information system, inventory management, Coordination between customer and company been identified as barriers in the reverse supply chain coordination mechanisms. These are shown in Figure.

Step 3: The DMs were asked to give the relative weight to each criterion according to the linguistic variable as per Table, (Tolga, Demircan, and Kahraman, 2005). After the criteria have been determined as given in Figure, a questionnaire has been prepared to determine the importance levels of these criteria. To evaluate the questions, experts only select the related linguistic variable. Further, for calculations they are converted into the corresponding TFNs.

Step 4: Fuzzy important weight of the criteria is calculated by taking average mean of the responses of the experts, this is shown in Table.

Step 5: Crisp relative important weight for identified criteria is calculated using the extent analysis method proposed by Chang (1996) as explained previously in this paper by equations number. The Fuzzy values of paired comparison are converted to crisp value via the Chang's extent analysis (1996) given by table 4 to table 9.

Table 1: Level 3, detail for barriers

C1	Environmental-Driven Legislation
C2	Provide Facilities Dissassembling
C3	Revenue Management
C4	Customer Awareness And Responsibiity
C5	Waste Management
C6	Forecasting Return
C7	Location Routing
C8	Network Structure Of Eol
C9	Integration Of Rl Activities Into Rsc
C10	Coordination Of Information System
C11	Inventory Management
C12	Coordination Between Cunsumer &Company

Table 2: Linguistic variables and triangular fuzzy numbers

Equal Importance (EI)	(0,0,0.1)
Moderate Importanance (MI)	(0,0.1,0.3)
Strong Importance (SI)	(0.1,0.3,0.5)
Very Strong Importance (VI)	(0.3,0.5,0.7)
Extreme Importance (EXI)	(0.5,0.7,0.9)
Intermediate Values B/W Two Adjacent Judgement (JU)	(0.7,0.9,1)
Must (MU)	(0.9,1,1)

CR (consistency ratio) 0.034673435 < 0.10

As need to check the consistency should be less than 0.1

As per above situation we are going in the direction and after this is weight of all creteria

So all creteria weights are defined as priotrise the alternatives so in next stage we need to find the critcal factors as per the alternatives by fuzzy topsis method and after this we see the rank of all alternatives.

Table 3: Criteria alternatives

CRETERIA	CW
C1	0.43760398
C2	0.14752839
C3	0.81446377
C4	0.63310862
C5	0.88238631
C6	0.57329769
C7	0.96984507
C8	0.65857717
C9	0.18533562
C10	0.70959777
C11	0.65223759
C12	0.33601801

Table 4: Decision maker's vs. rating for attributes

DECISION MAKERS	RATING FOR ATTRIBUTES											
DM	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
DM1	MU	VI	VI	MU	SI	MI	VI	MU	JU	EXI	SI	SI
DM2	JU	EXI	SI	MU	EXI	SI	EXI	JU	JU	MI	SI	MI
DM3	JU	VI	MI	MU	JU	SI	MI	MU	EXI	SI	EXI	MI
DM1	(0.9,1,1)	(0.3,0.5,0.7)	(0.3,0.5,0.7	(0.9,1,1)	(0.1,0.3,0.5)	(0,0.1,0.3)	(0.3,0.5,0.7)	(0.9,1,1)	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.1,0.3,0.5)	(0.1,0.3,0.5)
DM2	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.1,0.3,0.5	(0.9,1,1)	(0.5,0.7,0.9)	(0.1,0.3,0.5)	(0.5,0.7,0.9)	(0.7,0.9,1)	(0.7,0.9,1)	(0,0.1,0.3)	(0.1,0.3,0.5)	(0,0.1,0.3)
DM3	(0.7,0.9,1)	(0.3,0.5,0.7)	(0,0.1,0.3)	(0.9,1,1)	(0.7,0.9,1)	(0.1,0.3,0.5)	(0,0.1,0.3)	(0.9,1,1)	(0.5,0.7,0.9)	(0.1,0.3,0.5)	(0.5,0.7,0.9)	(0,0.1,0.3)
AVERAGE=	(0.7,0.9,1)	(0.4,0.6,0.7)	(0.1,0.3,0.5	(0.9,1,1)	(0.5,0.6,0.8)	(0.1,0.2,0.4	(0.3,0.4,	(0.8,0.9,1)	(0.6,0.8,0.9))	(0.2,0.4,0.6))	(0.2,0.4,0.6)	(0,0.2,0.4)

Table 5: linguistic triangular fuzzy numbers

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	1,1,1	(1.75,1.5,1	(7,3,1.4)	0.7,0.9,1	1.4,1.5,1.25	7,4.5,2.5	2.3,2.2,2.3	0.8,1,1	1.1,1.1,1.1	3.5,2.2,1.6	3.5,2.2,1.6	0,4.5,2.5
C2	0.5,0.6,0.7	1,1,1	4,2,1.4	0.9,0.6,0.7	0.8,1,1.75	4,3,1.75	1.3,1.5,2.3	0.5,0.6,0.7	0.6,0.7,0.7	2,1.5,1.1	2,1.5,1.6	0,3,1.75
C3	0.1,0.3,0.5	0.2,0.5,0.7	1,1,1	0.1,0.3,0.5	0.2,.5,.6.25	1,1.5,1.2	0.3,0.7,1.6	0.1,0.3,0.5	0.1,0.3,0.5	0.5,0.75,0.8	0.5,0.75,0.8	0,1.5,1.25
C4	1.2,1.1,1	2.2,1.6,1.4	9,3.3,2	1,1,1	1.8,1.6,1.25	9,5,2.5	3,2.5,3.3	1.1,1.1,1	1.5,1.25,1.1	4.5,2.5,1.6	2.5,1.6	0,5,2.5
C5	0.7,0.6,0.8	0.7,1,1.1	5,2,1.6	0.5,0.6,0.8	1,1,1	5,3,2	1.6,1.5,2.6	0.6,0.6,0.8	0.8,0.75,0.8	2.5,1.5,1.3	2.5,1.5,1.3	0,3,2
C6	0.1,0.2,0.4	0.25,0.3,0.	1,0.6,0.8	0.1,0.2,0.4	0.2,0.3,0.5	1,1,1	3.3,0.5,1.3	0.1,0.25,0.4	0.6,0.2,0.4	0.5,0.5,0.6	0.5,0.5,0.6	0,1,1
C7	0.4,0.4,0.3	0.7,0.6,0.4	0.3,1.3,0.6	0.3,0.4,0.3	0.6,0.6,0.3	3,2,0.7	1,1,1	0.3,0.4,0.3	2,0.5,0.3	1.5,1,2	1.5,1,2	0,2,0.7
C8	1.1,1,1	2,1.5,1.4	8,3,2	0.8,0.9,1	1.6,1.5,1.25	8,4.5,2.5	2.6,1.8,0.3	1,1,1	1.3,1.1,1.1	4,2.25,1.6	4,2.2,1.6	0,4.5,2.5
C9	0.8,0.8,0.9	1.5,1.3,1.2	6,2.6,1.8	0.6,1.3,0.9	1.2,1.3,1.1	6,4,2	2,2,3	0.75,0.8,3	1,1,1	3,2,1.6	3,2,1.5	0,4,2
C10	0.2,0.4,0.6	0.5,0.6,0.8	2,1.3,1.2	0.2,0.4,0.6	0.4,0.6,0.7	2,2,1.5	0.6,1,2	0.2,0.4,0.6	0.3,0.5,0.6	1,1,1	1,1,1.5	0,2,1.5
C11	0.2,0.5,0.6	0.5,0.6,0.8	2,1.3,1.2	0.2,0.4,0.6	0.4,0.6,0.7	0.2,1,1.5	0.4,1,2	0.2,0.5,0.6	0.6,0.5,0.6	1,1,1	1,1,1	0,2,1.5
C12	0,0.6,0.4	0,0.2,0.5	0,0.2,0.8	0,0.2,0.4	0,0.3,0.5	0,1,1	0,0.5,0.8	0,0.2,0.4	0,0.5,0.5	0,0.5,0.6	0,0.5,0.6	1,1,1

Table 6: linguistic rating by decision makers and solved fuzzy matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1	1	1.2	4	0.6	1.6	4.2	3.2	0.9	1.1	2.1	0.6	0.9
C2	0.6	1	2.3	0.7	1.1	2.75	1.5	0.6	0.6	1.5	1.2	2.7
C3	0.3	0.4	1	0.3	2	3.3	0.8	0.3	0.3	0.7	1.7	1.6
C4	1.1	1.6	5	1	1.6	5.2	2.6	1.1	1.2	2.6	0.5	0.6
C5	0.7	0.9	2.8	0.6	1	1	1.1	0.6	0.8	0.5	1.2	0.3
C6	0.2	0.5	0.8	0.2	0.3	1	1	0.2	0.4	1.5	1.5	0.9
C7	0.3	0.5	0.7	0.3	0.5	5.6	1	0.3	0.3	2.2	2.8	2.3
C8	1	1.6	4	0.9	1.6	4	2.1	1	1.1	2.3	2.2	2
C9	0.8	1	3.3	0.4	1.2	4	1.02	0.7	1	2.2	1.2	1.2
C10	0.4	0.6	1.5	0.4	0.5	1.3	1.2	0.4	0.4	1	1.2	1.2
C11	0.4	0.6	1.5	0.4	0.5	0.9	1.1	0.4	0.5	1	1	0.6
C12	0.3	0.2	0.3	0.2	0.2	0.7	0.4	0.2	0.3	0.3	0.3	1
SUM=	7.1	10.1	27.2	6	12.1	33.95	17.02	6.7	8	17.9	15.4	15.3

Table 7: Pairwise comparison of criteria via linguistic variables
TFN by decision makers average approach.

E	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	CRETERIA V
C1	0.14084507	0.11881	0.1470588	0.1	0.132231	0.12371134	0.1880141	0.134328358	0.1375	0.11731844	0.038961	0.0588235	0.437604
C2	0.08450704	0.09901	0.0845588	0.116666667	0.090909	0.08100147	0.08813161	0.089552239	0.075	0.08379888	0.0779221	0.1764706	0.147528
C3	0.04225352	0.0396	0.0367647	0.05	0.165289	0.09720177	0.04700353	0.044776119	0.0375	0.03910615	0.1103896	0.1045752	0.814464
C4	0.15492958	0.15842	0.1838235	0.166666667	0.132231	0.15316642	0.15276146	0.164179104	0.15	0.1452514	0.0324675	0.0392157	0.633109
C5	0.09859155	0.08911	0.1029412	0.1	0.082645	0.02945508	0.06462985	0.089552239	0.1	0.02793296	0.0779221	0.0196078	0.882386
C6	0.02816901	0.0495	0.0294118	0.033333333	0.024793	0.02945508	0.05875441	0.029850746	0.05	0.08379888	0.0974026	0.0588235	0.573298
C7	0.04225352	0.0495	0.0257353	0.05	0.041322	0.16494845	0.05875441	0.044776119	0.0375	0.12290503	0.1818182	0.1503268	0.969845
C8	0.14084507	0.15842	0.1470588	0.15	0.132231	0.11782032	0.12338425	0.149253731	0.1375	0.12849162	0.1428571	0.130719	0.658577
C9	0.11267606	0.09901	0.1213235	0.066666667	0.099174	0.11782032	0.05992949	0.104477612	0.125	0.12290503	0.0779221	0.0784314	0.185336
C10	0.05633803	0.05941	0.0551471	0.066666667	0.041322	0.03829161	0.07050529	0.059701493	0.05	0.05586592	0.0779221	0.0784314	0.709598
C11	0.05633803	0.05941	0.0551471	0.066666667	0.041322	0.02650957	0.06462985	0.059701493	0.0625	0.05586592	0.0649351	0.0392157	0.652238
C12	0.04225352	0.0198	0.0110294	0.033333333	0.016529	0.02061856	0.02350176	0.029850746	0.0375	0.01675978	0.0194805	0.0653595	0.336018
													1

Table 8: Normalised matrix

STEP :3 CALCULATING THE CONSISTENCY															
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	WEIGHTED SUM VALU	CRETERIA WEIGHT	λ_{max}
C1	0.43760398	0.52512	1.7504159	0.26256239	0.700166	1.83793673	1.40033275	0.393843585	0.481364382	0.91896837	0.2625624	0.3938436	9.364725251	17.29184652	0.541569
C2	0.08851704	0.14753	0.3393153	0.103269876	0.162281	0.40570309	0.22129259	0.088517037	0.088517037	0.22129259	0.1770341	0.3983267	2.441594931	3.794672825	0.643427
C3	0.24433913	0.32579	0.8144638	0.244339132	1.628928	2.68773046	0.65157102	0.244339132	0.244339132	0.57012464	1.3845884	1.303142	10.34368994	20.44304074	0.505976
C4	0.69641948	1.01297	3.1655431	0.633108618	1.012974	3.29216482	1.64608241	0.69641948	0.759730342	1.64608241	0.3165543	0.3798652	15.2579177	28.81941592	0.529432
C5	0.61767042	0.79415	2.4706817	0.529431788	0.882386	0.88238631	0.97062495	0.529431788	0.705909051	0.44119316	1.0588636	0.2647159	10.14744261	19.6772148	0.515695
C6	0.11465954	0.28665	0.4586382	0.114659539	0.171989	0.57329769	0.57329769	0.114659539	0.229319078	0.85994654	0.8599465	0.5159679	4.873030402	9.631401266	0.505952
C7	0.29095352	0.48492	0.6788915	0.29095352	0.484923	5.43113237	0.96984507	0.29095352	0.29095352	2.13365915	2.7155662	2.2306437	16.29339712	32.29584072	0.504505
C8	0.65857717	1.05372	2.6343087	0.59271945	1.053723	2.63430867	1.38301205	0.658577167	0.724434884	1.51472748	1.4488698	1.3171543	15.67413657	29.68969598	0.527932
C9	0.14826849	0.18534	0.6116075	0.074134246	0.222403	0.74134246	0.18904233	0.129734931	0.185335616	0.40773836	0.2224027	0.2224027	3.339747804	5.531227114	0.603799
C10	0.28383911	0.42576	1.0643966	0.283839107	0.354799	0.9224771	0.85151732	0.283839107	0.283839107	0.70959777	0.8515173	0.8515173	7.16693744	14.05003577	0.510101
C11	0.26089504	0.39134	0.9783564	0.260895038	0.326119	0.58701383	0.71746135	0.260895038	0.326118797	0.65223759	0.6522376	0.3913426	5.804914586	11.34893414	0.511494
C12	0.1008054	0.0672	0.1008054	0.067203602	0.067204	0.23521261	0.1344072	0.067203602	0.100805403	0.1008054	0.1008054	0.336018	1.478479247	2.856153092	0.517647
SUM=	3.94254832	5.7005	15.067424	3.457116307	7.067895	20.2307061	9.70848673	3.758413926	4.420666348	10.1763735	10.050948	8.6049399	102.1860136	195.4294789	6.417529

Table 9: Finding consistency of the weghted matrix

CONSISTENCY	0.507497394												
CR	N	1	2	3	4	5	6	7	8	9	10	11	12
	RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.54

5. Using Topsis

5.1 Prioritizing critical factor alternatives using Fuzzy TOPSIS technique

The objective of using Fuzzy TOPSIS technique is to find critical factors in RSCM based on identified coordination criteria for improving Reverse supply chain performance. After finding the important weights of the criteria, we will now use the Fuzzy TOPSIS technique to rank the reverse supply chain alternatives. To validate the proposed model, has been conducted to identify critical factors. Here, we have considered three most significant alternatives of the RSCM A1, A2, A3, A4, A5 and A6 TO RANK them of their attributes. The evaluation criteria environment-friendly, product acquisition, inventory reduction, product analysis, reverse distribution, cost minimization.

In the present case, a committee of three Decision makers DM1, DM2 and DM 3 is constituted to make their judgments for criteria based on SIX coordination alternatives environment-friendly, product acquisition, inventory reduction, product analysis, reverse distribution, cost minimization.

Then, the proposed methodology, which is integration of Fuzzy AHP and Fuzzy TOPSIS, is used to prioritize identified critical factor based on 12 coordination criteria, Environmental-Driven-Legislation, Provides Facilities, Dissassembling, Revenue, Management, Customer Awareness and -Responsibility, Waste Management, Forecasting Return, Location Routing, Network Structure of Eol, Integration of Rl Activities Into Rsc, Coordination of Information System, Inventory Management, Coordination Between Consumer.

Fig 10: Linguistic variables for rating.

Linguistic variables	**Triangular fuzzy numbers (TFN**
Equal Importance (EI)	(0,0,0.1)
Moderate Importanance (MI)	(0,0.1,0.3)
Essential Or Strong Importance (SI)	(0.1,0.3,0.5)
Very Strong Importance (VI)	(0.3,0.5,0.7)
Extreme Importance (EXI)	(0.5,0.7,0.9)
Intermediate Values B/W 2 Judgement (JU)	(0.7,0.9,1)
Must(MU)	(0.9,1,1)

The performance rating of three DMs on Reverse supply chain alternatives for each criterion, in linguistic variables are obtained and shown in Table 11**.**

Linguistic variables shown in Table 11 are converted into their corresponding TFNs, according to Table 12. This rating of the Reverse supply chain alternatives in TFNs is presented in Table 13. To form Fuzzy decision matrix, aggregated rating for the alternatives is calculated according to the steps suggested in algorithms of Fuzzy TOPSIS method. Table 14 report the Fuzzy rating of alternatives for each criterion.

Fuzzy decision matrix is normalized according to method suggested by Chen (2000) and shown in Table: Normalized Fuzzy decision matrix, with corresponding weight for each criterion is presented.

After normalization, a weighted normalized Fuzzy decision matrix is formed by multiplying the corresponding weight of each criterion as shown in Table: and weighted normalized Fuzzy decision matrix is shown in Table:

Table 11: Linguistic variables, to rate the alternatives with respect to each criterion.

DECISION MAKERS RATINGS									
CRITERION	ALTERNATIVES	DM1	DM2	DM3	CRITERION	ALTERNATIVES	DM1	DM2	DM3
C1	A1	MU	JU	MU	C7	A1	EXI	VI	EI
	A2	EXI	EXI	VI		A2	EXI	JU	JU
	A3	MI	SI	EI		A3	EXI	EXI	EXI
	A4	EI	MI	EI		A4	EXI	EXI	EXI
	A5	SI	SI	MI		A5	JU	JU	MU
	A6	MI	MI	MI		A6	JU	MU	MU
C2	A1	SI	SI	MI	C8	A1	JU	JU	JU
	A2	EXI	EXI	VI		A2	(0.9,1,1)	JU	MU
	A3	JU	JU	VI		A3	MU	MU	MU
	A4	MU	MU	JU		A4	JU	EXI	VI
	A5	VI	JU	EXI		A5	MU	JU	EXI
	A6	JU	MI	SI		A6	JU	EXI	EXI
C3	A1	MI	MI	MI	C9	A1	JU	MU	VI
	A2	JU	EXI	EXI		A2	EXI	JU	MU
ANY	A3	JU	JU	JU		A3	JU	MU	MU
	A4	MI	SI	SI		A4	JU	MU	MU
	A5	JU	MU	MU		A5	EXI	JU	EXI
	A6	MU	MU	MU		A6	JU	JU	EXI
C4	A1	MI	EI	SI	C10	A1	JU	EXI	EXI
	A2	MU	JU	JU		A2	MU	MU	JU
	A3	EI	EI	EI		A3	JU	EXI	VI
	A4	EXI	JU	VI		A4	EXI	VI	EXI
	A5	EI	EI	EI		A5	JU	EXI	EXI
	A6	SI	VI	SI		A6	EXI	EXI	JU
C5	A1	JU	JU	EXI	C11	A1	MI	VI	MI
	A2	EXI	EXI	VI		A2	EXI	VI	SI
	A3	JU	JU	EXI		A3	JU	JU	EXI
	A4	EXI	VI	EXI		A4	JU	EXI	JU
	A5	EI	EI	EI		A5	EI	EXI	VI
	A6	SI	VI	EI		A6	JU	EXI	VI
C6	A1	SI	VI	EXI	C12	A1	JU	EXI	VI
	A2	MU	MU	JU		A2	JU	JU	EXI
	A3	SI	VI	SI		A3	EI	EI	EI
	A4	EXI	EXI	EXI		A4	SI	VI	SI
	A5	EXI	EXI	EXI		A5	SI	SI	VI
	A6	SI	VI	VI		A6	JU	EXI	VI

Table 12: Fuzzy rating of alternatives for each criterion.

	A1	A2	A3	A4	A5	A6
C1	0.7,0.9,1	0.3,0.6,0.9	0,0.1,0.5	0,0.03,0.3	0,0.2,0.5	0,0.1,0.3
C2	0,0.2,0.5	0.3,0.6,0.9	0.3,0.7,0.9	0.7,0.9,1	0.3,0.7,1	0,0.4,1
C3	0,0.1,0.3	0.5,0.7,1	0.7,0.9,1	0,0.2,0.5	0.7,0.9,1	0.9,1,1
C4	0,0.1,0.5	0.7,0.9,1	0,0,0.1	0.3,0.7,1	0,0,0	0.1,0.3,0.7
C5	0.5,0.8,1	0.3,0.6,0.9	0.5,0.8,1	0.3,0.7,0.9	0,0,0.1	0,0.2,0.4
C6	0.1,0.5,0.9	0.7,0.9,1	0.1,0.3,0.5	0.5,0.7,0.9	0.5,0.7,0.9	0.1,0.4,0.6
C7	0,0.4,0.9	0.5,0.8,0.9	0.5,0.7,0.9	0.5,0.7,0.9	0.7,0.9,1	0.8,0.9,1
C8	0.7,0.9,1	0.7,0.9,1	0.9,1,1	0.3,0.7,1	0.5,0.8,1	0.5,0.7,1
C9	0.3,0.8,1	0.9,0.8,1	0.7,0.7,1	0.3,0.7,0.9	0.5,0.7,1	0.5,0.8,1
C10	0.5,0.8,1	0.7,0.9,1	0.3,0.7,1	0.3,0.6,0.9	0.5,0.7,1	0.5,0.8,1
C11	0,0.2,0.7	0.1,0.5,0.9	0.3,0.7,1	0.3,0.7,0.9	0.5,0.7,1	0.3,0.7,1
C12	0.3,0.7,1	0.5,0.8,1	0,0,0.1	0.1,0.3,0.7	0.1,0.3,0.7	0.3,0.7,1

Table 13: C

Table 14: Weighted fuzzy decision matrix

		DECISION MAKERS RATINGS							
CRETERION	ALTERNATIVES	DM1	DM2	DM3	CRITERION	ALTERNATIVES	DM1	DM2	DM3
C1	A1	(0.9,1,1)	(0.7,0.9,1)	(0.9,1,1)	C7	A1	(0.5,0.7,0.9)	(0.3,0.5,0.7)	(0,0,0.1)
	A2	(0.5,0.7,0.9	(0.5,0.7,0.9)	(0.3,0.5,0.7)		A2	(0.5,0.7,0.9)	(0.7,0.9,1)	(0.7,0.9,1)
	A3	(0,0.1,0.3)	(0.1,0.3,0.5)	(0,0,0.1)		A3	(0.5,0.7,0.9)	(0.5,0.7,0.9)	(0.5,0.7,0.9)
	A4	(0,0,0.1)	(0,0.1,0.3)	(0,0,0.1)		A4	(0.5,0.7,0.9)	(0.5,0.7,0.9)	(0.5,0.7,0.9)
	A5	(0.1,0.3,0.5	(0.1,0.3,0.5)	(0,0.1,0.3)		A5	(0.7,0.9,1)	(0.7,0.9,1)	(0.9,1,1)
	A6	(0,0.1,0.3)	(0,0.1,0.3)	(0,0.1,0.3)		A6	(0.7,0.9,1)	(0.9,1,1)	(0.9,1,1)
C2	A1	(0.1,0.3,0.5	(0.1,0.3,0.5)	(0,0.1,0.3)	C8	A1	(0.7,0.9,1)	(0.7,0.9,1)	(0.7,0.9,1)
	A2	(0.5,0.7,0.9	(0.5,0.7,0.9)	(0.3,0.5,0.7)		A2	(0.9,1,1)	(0.7,0.9,1)	(0.9,1,1)
	A3	(0.7,0.9,1)	(0.7,0.9,1)	(0.3,0.5,0.7)		A3	(0.9,1,1)	(0.9,1,1)	(0.9,1,1)
	A4	(0.9,1,1)	(0.9,1,1)	(0.7,0.9,1)		A4	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)
	A5	(0.3,0.5,0.7	(0.7,0.9,1)	(0.5,0.7,0.9)		A5	(0.9,1,1)	(0.7,0.9,1)	(0.5,0.7,0.9)
	A6	(0.7,0.9,1)	(0,0.1,0.3)	(0.1,0.3,0.5)		A6	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.5,0.7,0.9)
C3	A1	(0,0.1,0.3)	(0,0.1,0.3)	(0,0.1,0.3)	C9	A1	(0.7,0.9,1)	(0.9,1,1)	(0.3,0.5,0.7)
	A2	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.5,0.7,0.9)		A2	(0.5,0.7,0.9)	(0.7,0.9,1)	(0.9,1,1)
	A3	(0.7,0.9,1)	(0.7,0.9,1)	(0.7,0.9,1)		A3	(0.7,0.9,1)	(0.9,1,1)	(0.9,1,1)
	A4	(0,0.1,0.3)	(0.1,0.3,0.5)	(0.1,0.3,0.5)		A4	(0.7,0.9,1)	(0.9,1,1)	(0.9,1,1)
	A5	(0.7,0.9,1)	(0.9,1,1)	(0.9,1,1)		A5	(0.5,0.7,0.9)	(0.7,0.9,1)	(0.5,0.7,0.9)
	A6	(0.9,1,1)	(0.9,1,1)	(0.9,1,1)		A6	(0.7,0.9,1)	(0.7,0.9,1)	(0.5,0.7,0.9)
C4	A1	(0,0.1,0.3)	(0,0,0.1)	(0.1,0.3,0.5)	C10	A1	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.5,0.7,0.9)
	A2	(0.9,1,1)	(0.7,0.9,1)	(0.7,0.9,1)		A2	(0.9,1,1)	(0.9,1,1)	(0.7,0.9,1)
	A3	(0,0,0.1)	(0,0,0.1)	(0,0,0.1)		A3	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)
	A4	(0.5,0.7,0.9	(0.7,0.9,1)	(0.3,0.5,0.7)		A4	(0.5,0.7,0.9)	(0.3,0.5,0.7)	(0.5,0.7,0.9)
	A5	(0,0,0.1)	(0,0,0.1)	(0,0,0.1)		A5	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.5,0.7,0.9)
	A6	(0.1,0.3,0.5	(0.3,0.5,0.7)	(0.1,0.3,0.5)		A6	(0.5,0.7,0.9)	(0.5,0.7,0.9)	(0.7,0.9,1)
C5	A1	(0.7,0.9,1)	(0.7,0.9,1)	(0.5,0.7,0.9)	C11	A1	(0,0.1,0.3)	(0.3,0.5,0.7)	(0,0.1,0.3)
	A2	(0.5,0.7,0.9	(0.5,0.7,0.9)	(0.3,0.5,0.7)		A2	(0.5,0.7,0.9)	(0.3,0.5,0.7)	(0.1,0.3,0.5)
	A3	(0.7,0.9,1)	(0.7,0.9,1)	(0.5,0.7,0.9)		A3	(0.7,0.9,1)	(0.7,0.9,1)	(0.5,0.7,0.9)
	A4	(0.5,0.7,0.9	(0.3,0.5,0.7)	(0.5,0.7,0.9)		A4	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.7,0.9,1)
	A5	(0,0,0.1)	(0,0,0.1)	(0,0,0.1)		A5	(0,0,0.1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)
	A6	(0.1,0.3,0.5	(0.3,0.5,0.7)	(0,0,0.1)		A6	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)
C6	A1	(0.1,0.3,0.5	(0.3,0.5,0.7)	(0.5,0.7,0.9)	C12	A1	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)
	A2	(0.9,1,1)	(0.9,1,1)	(0.7,0.9,1)		A2	(0.7,0.9,1)	(0.7,0.9,1)	(0.5,0.7,0.9)
	A3	(0.1,0.3,0.5	(0.3,0.5,0.7)	(0.1,0.3,0.5)		A3	(0,0,0.1)	(0,0,0.1)	(0,0,0.1)
	A4	(0.5,0.7,0.9	(0.5,0.7,0.9)	(0.5,0.7,0.9)		A4	(0.1,0.3,0.5)	(0.3,0.5,0.7)	(0.1,0.3,0.5)
	A5	(0.5,0.7,0.9	(0.5,0.7,0.9)	(0.5,0.7,0.9)		A5	(0.1,0.3,0.5)	(0.1,0.3,0.5)	(0.3,0.5,0.7)
	A6	(0.1,0.3,0.5	(0.3,0.5,0.7)	(0.3,0.5,0.7)		A6	(0.7,0.9,1)	(0.5,0.7,0.9)	(0.3,0.5,0.7)

CW		A1	A2	A3	A4	A5	A6
0.437604	C1	0.2,0.3,0.4	0.1,0.2,0.3	0,0.04,0.2	0,0.1,0.1	0,0.8,0.2	0,0.04,0.1
0.147528	C2	0,0.02,0.03	0.03,0.06,0.0	0.03,0.07,0.09	0.07,0.09,0	0.03,0.07,0.1	0,0.04,0.1
0.814464	C3	0,0.08,0.23	0.4,0.5,0.8	0.56,0.7,0.8	0,0.1,0.4	0.5,0.7,0.8	0.7,0.8,0.8
0.633109	C4	0,0.06,0.3	0.4,0.5,0.6	0,0,0.06	0.1,0.7,1	0,0,0	0.1,0.3,0.7
0.882386	C5	0.5,0.8,0.8	0.3,0.6,0.9	0.5,0.8,1	0.3,0.4,0.5	0,0,0.06	0,0.1,0.2
0.573298	C6	0.05,0.2,0.	0.3,0.4,0.5	0.05,0.1,0.2	0.2,0.3,0.4	0.2,0.3,0.4	0.05,0.2,0.3
0.969845	C7	0,0.3,0.8	0.4,0.7,0.8	0.4,0.6,0.8	0.4,0.6,0.8	0.6,0.8,0.9	0.7,0.8,0.9
0.658577	C8	0.4,0.5,0.6	0.4,0.5,0.6	0.5,0.6,0.6	0.01,0.4,0.6	0.3,0.4,,6	0.3,0.4,0.6
0.185336	C9	0.03,0.08,0	0.09,0.08,0.1	0.07,0.07,0.1	0.03,0.07,0	0.05,0.07,0.1	0.05,0.08,0.1
0.709598	C10	0.3,0.5,0.7	0.4,0.6,0.7	0.2,0.4,0.7	0.2,0.4,0.6	0.3,0.4,0.7	0.3,0.5,0.7
0.652238	C11	0,0.1,0.4	0.06,0.3,0.5	0.1,0.4,0.6	0.1,0.4,0.6	0.3,0.4,0.6	0.1,0.4,0.6
0.336018	C12	0.9,0.2,0.3	0.1,0.2,0.3	0,0,0.3	0.03,0.09,0	0.03,0.09,0.2	0.9,0.2,0.3

P=(0,0,0,0,0,0,0,0,0,0,0,),(0.03,0.030,0.03,0.03,0.03,0.03,0.0

3,0.03,0.030.030.03,0.3)(0,0,0,0,0,0,0,0,0,0,0,)(0,0,0,0,0,0,0,0,0,0,0,)(0,0,0,0,0,0,0,0,0,0,0,)(0,0,0,0,0,0,0,0,0,0,0,)

P*=(0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,),(0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,)(0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,

Closeness Coefficient of Alterntives and Ranking					
	di*	di-	di* + di-	Cci= di-\di+di*	Rank
A1	3.16	2.98	6.14	0.48	1
A2	3.9	2.3	6.2	0.387	5
A3	5.59	3.99	9.58	0.416	2
A4	6	4.02	10.02	0.401	4
A5	5.22	2.8	8.02	0.34	6
A6	6.4	4.47	10.87	0.411	3

0.8,0.8,)(0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,0.8,)(0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,)(0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9,0.9)

Now, both Fuzzy positive ideal solution (FPIS) and Fuzzy negative ideal solution (FNIS) are calculated after determining FPIS and FNIS, the distances of each alternative from FPIS and FNIS with respect to each criterion are calculated using Vertex method (Chen, 2000) as

$$d_1^* = d(P_1, P^*)$$

$$d_1^- = d(P_1, P^-)$$

Table 15: Distance between Pi (I=1, 2, 3...6) and P* w.r.t each criterion (C1, C2, C3, C4……C12)

Creteria Weight	Creteria	D1*	d2*	d3*	d4*	d5*	6*
CW		A1	A2	A3	A4	A5	A6
0.43760398	C1	0.03	0.4	0.7	0.7	0.06	0.8
0.14752839	C2	0.06	0.2	0.7	0.7	0.7	0.9
0.81446377	C3	0.06	0.3	0.2	0.6	0.2	0.4
0.63310862	C4	0.03	0.4	0.7	0.4	0.9	0.5
0.88238631	C5	0.02	0.3	0.2	0.4	0.9	0.6
0.57329769	C6	0.06	0.4	0.6	0.5	0.6	0.4
0.96984507	C7	0.5	0.3	0.2	0.3	0.1	0.2
0.65857717	C8	0.4	0.4	0.2	0.5	0.06	0.4
0.18533562	C9	0.4	0.1	0.59	0.7	0.3	0.6
0.70959777	C10	0.4	0.4	0.4	0.4	0.3	0.7
0.65223759	C11	0.7	0.3	0.4	0.5	0.4	0.5
0.33601801	C12	0.5	0.4	0.7	0.3	0.7	0.4
	SUM	3.16	3.9	5.59	6	5.22	6.4

Table 16: Distance between Pi (i=1, 2, 3...6) and P w.r.t each criterion (C1, C2, C3, C4..C12).

		d1-	d2-	d3-	d4-	d5-	d6-
CW		A1	A2	A3	A4	A5	A6
0.43760398	C1	0.3	0.4	0.2	0.4	0.7	0.2
0.14752839	C2	0.09	0.3	0.4	0.3	0.02	0.2
0.81446377	C3	0.21	0.03	0.3	0.2	0.5	0.7
0.63310862	C4	0.02	0.2	0.05	0.05	0.2	0.4
0.88238631	C5	0.7	0.1	0.4	0.4	0	0.02
0.57329769	C6	0.03	0.1	0.6	0.6	0.1	0.03
0.96984507	C7	0.4	0.2	0.5	0.7	0.2	0.8
0.65857717	C8	0.03	0.01	0.7	0.5	0.6	0.4
0.18533562	C9	0.1	0.3	0.04	0.2	0.4	0.02
0.70959777	C10	0.4	0.02	0.4	0.07	0.02	0.6
0.65223759	C11	0.2	0.04	0.2	0.2	0.03	0.4
0.33601801	C12	0.5	0.6	0.2	0.4	0.03	0.7
	sum	2.98	2.3	3.99	4.02	2.8	4.47

Table 17: Closeness coefficient of alternatives and their ranking

Similarly, other values of di* for six alternatives with respect to each criterion have been calculated. These values are shown in Table. Further, values of di− for six alternatives with respect to each criterion have been calculated. These values are shown. Then, the closeness coefficient of six alternatives CCi (i=1, 2, 3…..) are calculated by using and results are presented in Table.

$$CC_i = \frac{d_i^-}{d_i^- + d_i^*}$$

The closeness coefficient represents the distance to the FPIS (P*) and Fuzzy negative ideal solution (P−). Table provides the values of closeness coefficients for all alternatives. The value of closeness coefficient for first alternative is 0.48, for second 0.416, and for third is 0.411, 0.401, 0.387, and 0.34**).** This indicates that first alternative is closest to FPIS and should be given first critical factors. Based on closeness coefficients given in the Table, the ranking order of alternatives have been determined, the priorities of the alternatives are A1 >A3 >A6>A4>A2>A5. The first alternatives determined as most appropriate critical factor for the RSCM under consideration, because the first alternatives is closer to the FPIS and farther from the FNIS. Similarly, 5th alternatives for which closeness coefficient value is lowest, identified as the least preferred because it farther to FPIS and close to FNIS [13].

5. Conclusions

In this paper, we worked on various peer-reviewed papers for identifying the barriers while implementing reverse supply chain management in any organization and after that find its critical factors using multi-criteria decision-making model which has been developed for identifying the critical factors in reverse supply chain management. This approach is used in various research paper after checking all the suitability of vagueness associated with the independent perception of decision-makers which has been applied. The Role of the decision makers is to prioritization of the factor's crucial strategic way for survival of the firm, because the profitability of a firm and environmental satisfaction is directly proportional to the effectiveness of the prioritization process on their judgment level. It has been detected from the various literature that decision-makers face the uncertainties from subjective perceptions and experiences in the decision-making process. We take help from decision makers who are expert in their field and calculate their process alternatives Using Fuzzy AHP and Fuzzy TOPSIS, uncertainty and vagueness can be effectively handled and reached efficiently. As a result, Fuzzy AHP helps in identifying consistent information between the decision

makers is determined as the most important criterion for coordination, because this criterion has highest weight priority. Then, the alternatives are ranked based on closeness coefficient level, which is calculated for each partner using Fuzzy TOPSIS. It has been distinguished from the different literary works that chiefs face the vulnerabilities from abstract observations and encounters in the dynamic procedure. Utilizing Fuzzy AHP and Fuzzy TOPSIS, vulnerability and dubiousness can be viably taken care of and reached effectively. Accordingly, Fuzzy AHP helps in recognizing steady Information exchanges between the chiefs is resolved as the most significant rule for coordination, since this measure has most elevated weight need. At that point, the options are positioned depends upon the coefficient of closeness, that is determined for each accomplice utilizing Fuzzy TOPSIS.

Alternatives 'A4' is the most appropriate accomplice for the task is viewed as most basic components. The fourth option is resolved as most fitting option for the Reverse flexibly chain viable, in light of the fact that the fourth option is closer to FPIS than the FNIS as shown in the results. So also, for that coefficient of esteem is most reduced identified third option the least liked, in light of the fact that it is farther to FPIS and near FNIS. The concept is valuable in tackling the commonsense issue, since ambiguity and imprecision can be successfully dealt with in this model. On the off chance that the measures and options are obviously defined, the current model can be received in any industry.

References

[1]. Ttp://home.iitk.ac.in/~pmehta/scmqip07/reverse_logistics.ppt
[2]. https://www.cognizant.com/whitepapers/reverse-supply-chain.pdf
[3]. http://www.divaportal.org/smash/get/diva2:321961/fulltext01.pdf
[4]. A Agrawal, RK Singh, K Rajesh, Q Murtaza. Reverse supply chain issues in indian electronics industry: a case study, journal of remanufacturing, , heidelberg, 8(3), 2018, 115-129, http://dx.doi.org/10.1007/s13243-018-0049-7
[5]. http://home.nitk.ac.in/~vsharma/reverse logistics.ppt
[6]. A Agrawal, RK Singh, K Rajesh, Q Murtaza. Prioritizing critical success factors for reverse logistics implementation using fuzzy-tops is methodology, J Ind Eng. Int 12, 2016, 15–27
[7]. A new car selection in the market using topsis technique international journal of engineering research and general science 2(4,) 2014
[8]. S Agrawal, RK Singh, Q Murtaza. Forecasting product returns for recycling in indian electronics industry. J adv. manage res 11(1), 2014,102–114
[9]. S Agrawal, RK Singh, Q Murtaza. A literature review and perspectives in reverse logistics. Resource conservation recycle 97, 2015, 76–92 3.
[10]. S Agrawal, RK Singh, Q Murtaza. Outsourcing decisions in reverse logistics: sustainable balanced scorecard and graph theoretic approach. Resource conservation recycle 108, 2016, 41–53
[11]. S Agrawal, RK Singh, Q Murtaza. Triple bottom line performance evaluation of reverse logistics. Compet rev 26(3):289–310 6.
[12]. PS Aithal. A review on opportunities and challenges for mobile business activities in India. Int J Manage It Eng 6(1), 2016, 124–148
[13]. A Agarwal, R Shankar. On-line trust building in e-enabled supply chain. Supply Chain Management: An International Journal, 8, 2003, 324–334
[14]. Li Chunhui, Li Aizhen. The Application of TOPSIS Method to Comprehensive Assessment of Environmental Quality, Journal of Geological Hazard and Environmental Preservation, 10(2), 1999, 9 – 13.
[15]. [21] arshinder, n. A., kanda, a., & deshmukh, s. G. (2006). A coordination-based perspective on the procurement process in the supply chain. International journal of value chain management,
[16]. [22]. Loo-see beh abby ghobadian qile he david gallear nicholas o'regan , (2016),"second-life retailing: a reverse supply chain perspective", supply chain management: an international journal, vol. 21 iss 2 pp. Permanent link to this document: http://dx.doi.org/10.1108/scm-07-2015-0296
[17]. [23]. Mohd rizaimy shaharudin, kannan govindan, suhaiza zailani, keah choon tan, mohammad iranmanesh, product return management: linking product returns, closed-loop supply chain activities and the effectiveness of the reverse supply chains, journal (2017), doi: 10.1016/j.jclepro.2017.02.133
[18]. [24].rogers, ds & tibben-lembke, an examination of reverse logistics practices', journal of business logistics, vol.22, no. 2, 2001, pp.129-48.
[19]. [25] bayles, bayles, 'return merchandise authorization process, send it back! The role of reverse logistics', no. 5 december 2003, http://www.informit.com/articles/article.asp?p=164926&seqnum7.
[20]. [27].tahaz,rostams. Afuzzyahp–ann-based decision support system for machine tool selection in a flexible manufacturing cell .the international journal of advanced manufacturing technology.2011; 57(5–8):719–733.
[21]. [28]hindawaji advances in fuzzy systems volumn 2018,article id: 14pages https:doi.org/10.1155/2018/6703798
[22]. [29] matopoulos, a., vlachopoulou, m., manthou, v. “understanding the factors affecting e-business adoption and impact on logistics processes”, journal of manufacturing technology management, vol. 20, no. 6, pp. 853-865, 2009.
[23]. [30]wu, j. Z., hsu, c. Y. “critical success factors for improving decision quality on collaborative design in the ic supply chain”, journal of quality, vol. 16, no. 2, pp. 95-108, 2009.
[24]. [31] al-fawaz, k., al-salti, z., eldabi, t. “critical success factors in erp implementation: a review” european and mediterranean conference on information systems 2008 (emcis2008), dubai, pp. 1-9, 2008.
[25]. [32]soin, s. S. “critical success factors in supply chain management at high technology companies” dba thesis, university of southern queensland, 2004.[32]soin, s. S. “critical success factors in supply chain management at high technology companies” dba thesis, university of southern queensland, 2004
[26]. [33]. Adeyeye, k., osmani, m. And brown, c. (2007), “energy conservation and building design: the environmental legislation push and pull factors”, structural survey, vol. 25 no. 5, pp. 375-90.
[27]. [34]. Dekker, r., fleischmann, m., inderfurth, k. And van wassenhove, l.n. (eds) (2003), reverse logistics: quantitative models for closed-loop supply chains, springer verlag, berlin
[28]. [35]. Adenso-diaz, b., moure, f. And rendueles, m., 2002. Automatic disassembly plans: applications to the continuous process industries. Journal of manufacturing systems, 21 (4), 276–286
[29]. [36].autry, c.w., daugherty, p.j. And richey, r.g., 2001. The challenge of reverse logistics in catalog retailing. International journal of physical distribution & logistics management, 31 (1), 26–37.
[30]. [37]. Ahluwalia, p.k. And nema, a.k., 2006. Multiobjective reverse logistics model for integrated computer waste management. Waste management and research, 24 (6), 514–527.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Impact of Supervised Classification Techniques for the Prediction of Student's Performance

Rahul
Department of CSE
Delhi Technological University
Delhi, India
rahul@dtu.ac.in

Rahul Katarya
Department of CSE
Delhi Technological University
Delhi, India
rahulkatarya@dtu.ac.in

Abstract - Every country's concern about its growth or development is education. This field creates a way to discover hidden examples from instructive information. The authors have researched by comparing the different classification techniques on the collected dataset which is present online on the UCI Machine Learning (ML) repository. The results of this learning identify an explanatory structure uniting multiple dimensions persuading the prediction. For this research, the authors conducted the experiments on the collected dataset using the Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) and measure the performance using the metrics like Accuracy (Acc.), Precision (Pr.) and Recall (Rec.). This research will also help the schools, colleges and university teachers or faculty for identifying the weak students in the class and to help them separately by conducting remedial classes or any other suitable method.

Keywords— Educational Data Mining (EDM), Machine Learning, Classification, Accuracy

I. INTRODUCTION

In a highly digital world, it is becoming increasingly important to recognize the skills of students early and guide them into suitable career paths [1]. Universities or colleges or schools today, which are almost identical to business organizations, work in a highly competitive and diverse world [2]. Every year, colleges, Universities, Education Centers, and several institutes accept students in different fields with specific grades and the capacity to succeed. Currently, in education areas for performance research and instruction improvement, Data Mining (DM), and other methods are becoming well known [2]. DM is described as a process that uses one or more software to process huge amounts of data to find the masked patterns in large data lots [3].

EDM is an independent field of study in the mining of data. EDM reviews data generated by the educational institution like a prediction of students results, learning analytics, a grouping of a student by performance and recommendation to the user [4]. Data can be something that is connected to society, student academic data and people's interests. From time to time the authors can see the new data emerging [5]. The techniques removed from EDM regulation are therefore used to comprehend understudies which include their learning conduct just as to foresee their scholarly presentation using various student-related attributes like demographic, academic, and external details. To maintain the excellence of education and to encourage skills-oriented research, the proper use of laboratory infrastructures and improve the quality of the placements which needs to be analyzed [6].

Much DM and ML technology are currently being used in the field of education to support both students and teachers for better results in teaching and learning [7]. The ML algorithms or techniques can be applied to students data to predict their success and maintain a good education [6]. This work is focused on the application of specific ML algorithms to data sets for students. This research would provide a more efficient analysis in which teachers or administrators do not only predict the failure of students but they can also divide them as per their level [8]. This approach ensures that every student has appropriate and necessary facilities so that they do not have any problems in their education [9]. Comparison is made of the supervised classification algorithms, comparing the efficiency of different models. The motive of this paper is to check out the shortcomings and abilities of the students from past experiences as well as helping to update or amend the curriculum with the aid of these predicted outcomes for the time table committees [3]. Success forecasts can also help admission committees figure out which applicants are most suitable and least suitable for an academic program [3]. The acts for poor students will be to reschedule remedial courses and give them extra assignments to enhance their results. This will also help decision-makers at universities as they set standards for efficient admissions [10]. Several innovations have been implemented, tested and are being developed to solve these security issues [11].

The structure of this paper can be categorized as Section 2 is about the related study in this area for predicting the POS. In Section 3, there is a discussion of ML techniques used in this paper for classification. Section 4 will have the methodology in which there is a discussion of the dataset collection and experimental results. In the end, this paper will have a conclusion and the future work for the researchers to improve the prediction while evaluating the POS.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

II. LITERATURE REVIEW

Abid Jamil et al. [12] proposed a technique in which they applied data pre-processing to pick the best attributes to improve their accuracy, then they applied various algorithms to test their accuracy and found that the RF algorithm yielded the best results by collecting online data set named as Aj Jaraha and they implemented different DM techniques like Naïve Bayes (NB), RF, J48, REP Tree and Zero R. Evawaty Tanuar et al. [7] implemented ML techniques like Generalized Linear Model, Deep Learning and DT and did the experiments on Grade Point Average (GPA) and a group of GPA and from these DM experiments using Rapid Miner on the result of the 1st Semester to predict the final year GPA, it has been shown that GPA grouping is stronger in runtime compared to the other data form. Fergie Joanda Kaunang et al. [13] implemented classification methods DT and RF using K-fold Cross Validation on the collected data using questionnaires and checked performance using evaluation metrics like Acc., Pr., Rec. and F- Measure.

Iti Burman et al. [14] used the SVM multi-classifier for classifying the students into the high, average and low categories by their academic scores and checked performance using evaluation metrics like Accuracy, Sensitivity and Specificity. It is done with Linear Kernel and Radial Base Kernel. Mehdi Mohammadi et al. [2] implemented supervised learning algorithms like KNN, DT and NB and checked accuracy on all these algorithms. They used KNN, DT and NB Classifiers on the dataset of 230 Kabul University students to predict their GPA as high, medium and low. Nikola et al. [15] were intended to provide overview and contrast between supervised ML approaches used for student success prediction i.e. classifying high-risk students to drop out of the course and predicting their potential accomplishments or final exam scores. Raza Hasan et al. [16] implemented the classification algorithms like J48, REP Tree, RF, Logistic Model Tree, Hoeffding Tree, Decision Stump, NB and SMO and found the Accuracy and Kappa Value on these algorithms. Their proposed work uses the Electronic Commerce Technology Platform as it is a central module provided in every specialization of computing. Rojina Deuza et al. [17] determined the accuracy between the different DT Algorithms like J48, Decision Table, One R, Decision Stump and PART and also implemented Multilayer Perceptron (MLP) using neural network (NN) and compared accuracy with hidden layer neurons.

Vahide Nida Uzel et al. [18] implemented the various classification algorithms like MLP, RF, NB, J48 and Voting Classifiers and checked performance using evaluation metrics like Accuracy, Precision, Recall and F- Measure. They also observed which features are more related to the students' progress in academic performance. Yasmeen Shaher Alsalman et al. [19] implemented the classification techniques J48 and Artificial Neural Network (ANN) on the dataset collected using Questionnaire and certain attributes were selected to assess their relevance to a Jordanian University academic POS. Abinav et al. [8] suggested better ML methods in terms of Acc. and other attributes and parameters while choosing. They used pre-processing data, extraction of features, fine-tuning of algorithm parameters to construct a model that is not targeted by many researchers. Zhengyang et al. [20] proposed a novel approach to Exam paper generation (EPG), using the Deep Knowledge Tracing (DKT) model to predict learning results, and using dynamic programming and genetic algorithms to improve the quality of examination papers. With the higher precision of student performance estimation, their methodology will help ensure the examination paper's complexity and unbiased scores distribution between the students obtained in the test.

III. ML TECHNIQUES USED FOR CLASSIFICATION

In this section, there are four classification techniques which are used in this research and will discuss in brief one by one as follows:

A. Decision Tree

DTs are the most well-known classification model. It is a tree where the node of each branch is an option between many alternatives and the node of each leaf is a decision. The DT starts with the root node for users to take action. Through this node, users divide each node recurrently by DT algorithm of learning. The result is a DT with each branch reflecting a potential decision scenario and its outcome [4].

B. Random Forest

It is another DT induction algorithm. Rather than learning a DT, RF aims to increase the rate of classification accuracy by studying and using more than one DT [18]. RF classifier creates a set of DTs from randomly chosen training set subsets. The final class of the test object is determined by the aggregation of votes from various DTs [8].

C. K-Nearest Neighbour

This algorithm reflects a different approach to classification. KNN is known to be an algorithm of un parametrically supervised learning. Here the samples are categorized based on the nearest training samples present within the feature space [9]. They don't create any simple universal model but just estimate it locally and implicitly. For KNN classifiers, the only learning function is to pick two essential parameters: the number of neighbours k and the distance metric d [4].

D. Support Vector Machine

It is a technique which is used to evaluate data classification. By finding the hyperplane, this algorithm will work by finding hyperplane that aims at widening the boundaries between two class sets. SVM focuses on the limits of a class only; points that are only classified quickly way are missed [4]. This algorithm uses the hyperplane to divide the dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

into groups, with the distance known as margin as wide as possible [14].

IV. METHODOLOGY

In this part, the authors will focus on the dataset collected and on the experimental results and analysis. This will be as follows:

A. Collection of Data

The dataset is available online, with information about the student of two Portuguese schools which have mathematics marks and Portuguese language marks. In this paper, the authors have only used the dataset of Portuguese. The collected data is given in the form of a .csv file and having 649 rows. It also includes 33 features but the authors have used only 16 features. The target variable is G3, which is the final result of the subject of Portuguese language whereas the G1 and G2 are the first and second-period grades respectively [21].

Table1: Name of Attributes used and their description [21]

Attributes	Description
sex	Sex of student
address	Home Address type
studytime	Weekly study time
failures	Number of past class failures
schoolsup	Extra Educational Support
famsup	Family Educational Support
higher	Wants to take higher education
Internet	Internet access at home
famrel	Quality of family relationships
goout	Going out with friends
Dalc	Workday alcohol consumption
health	Current health status
absences	Number of School absences
G1	First period grade
G2	Second period grade
G3	Final grade

B. Data Pre-processing

First, the authors collected the data and then they need to pre-process the data to remove any missing values or duplicates and to clean the data. They have to fill the missing values by using the method of mean and median or by using the most frequent values used in the dataset. At last, they divided the dataset into two parts: Training dataset and Testing dataset into 80% and 20% respectively.

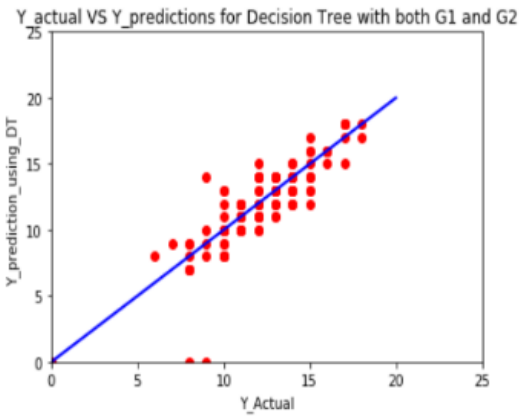
C. Experimental Results and Discussions

The authors conducted this research on the Jupyter Notebook using the Python libraries for ML in order the predict the performance of the students. They have used three different performance metrics are Accuracy, Precision and Recall for this research. Accuracy is the ratio of the total number of instances correctly estimated to the actual number of instances. Precision is the proportion of correctly predicted positive instances to the total positive instances expected. Recall is the proportion of correctly expected positive instances in the actual class to all instances [13]. They also visualised the results using matplotlib python library for better visualisation of the results. The POS calculated on the basis of classification algorithms and there are four different cases have taken for the evaluation i.e. G1 and G2 known, Only G1 Known, Only G2 Known, and Without G1 and G2 and these cases have taken for all 3 performance metrics (Acc., Pr. and Rec.). Now the authors showing all the cases for accuracy calculation in Table I are as follows:

TABLE I: CASES TAKEN FROM THE ATTRIBUTES FOR THE CALCULATION OF ACCURACY

S.No	Classification Algorithms	G1 & G2 Known	Only G1 Known	Only G2 Known	Without G1 and G2
1.	DT	0.3845	0.2537	0.3537	0.1537
2.	RF	0.4231	0.2422	0.4077	0.1614
3.	KNN	0.3154	0.1962	0.3037	0.1230
4.	SVM	0.4270	0.2641	0.3808	0.1577

From Table I, the authors found the best accuracy value for SVM in both the cases when G1 and G2 are known and when only G1 known. But when only G2 known and also without G1 and G2, found that the RF gave the best accuracy. The authors also visualized the prediction between the actual value and predicted value in Fig.1 by using the matplotlib library in python for the DT algorithm. In part a) of Fig.1, there is a visualization between the actual value (AV) on the x-axis and the predicted value (PV) on the y-axis with the case of having both G1 and G2 using DT algorithm. In part b) of Fig.1, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having the only G1 using the DT algorithm. In part c) of Fig.1, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having the only G2 using the DT algorithm. In part d) of Fig.1, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having without G1 and G2 using the DT algorithm.



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

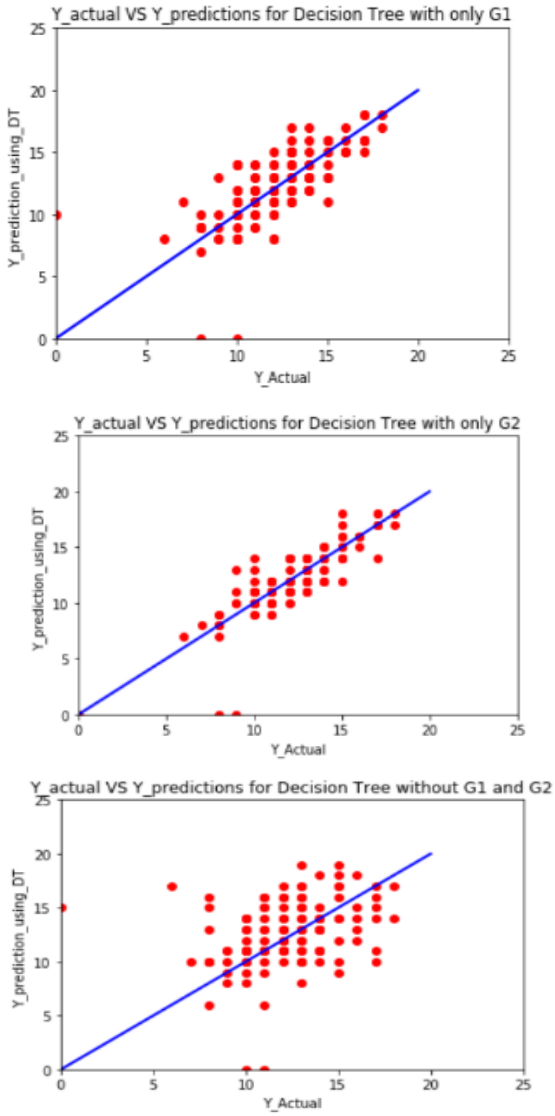


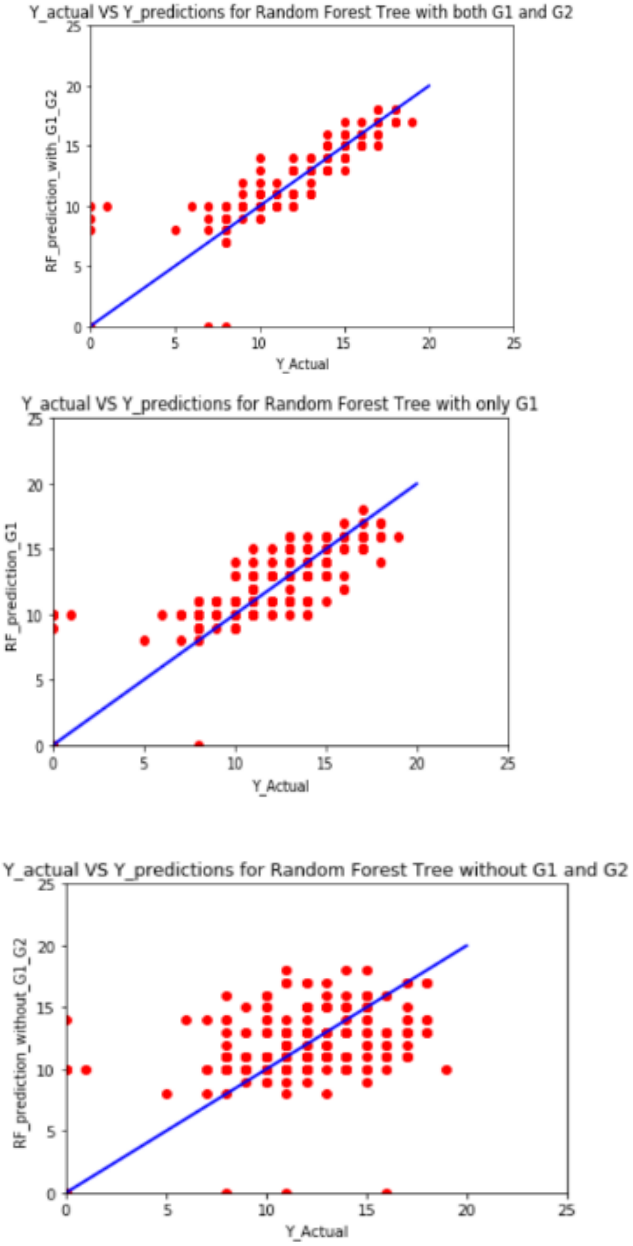
Fig. 1. Visualisation of the Actual Value vs Predicted Value using matplotlib library a) Both G1 and G2 known b) Only G1 Known c) Only G2 Known d) Without G1 and G2 by using DT

Next, the authors have calculated the Precision for all the cases to evaluate performance prediction by applying the classification algorithms in Table II are as follows:

TABLE II: CASES TAKEN FROM THE ATTRIBUTES FOR THE CALCULATION OF PRECISION VALUE

S.No	Classification Algorithms	G1 & G2 Known	Only G1 Known	Only G2 Known	Without G1 and G2
1.	DT	0.39	0.27	0.38	0.15
2.	RF	0.37	0.22	0.34	0.15
3.	KNN	0.28	0.17	0.28	0.09
4.	SVM	0.41	0.25	0.37	0.10

From Table II, the authors found the best precision value for SVM in the case when G1 and G2 are known but when only G1 known and only G2 known, the DT gave the best precision value in both the cases. The authors also visualized the prediction between the actual value and predicted value in Fig.2 by using the matplotlib library in python for the RF algorithm. In part a) of Fig.2, there is a visualization between the actual value (AV) on the x-axis and the predicted value (PV) on the y-axis with the case of having both G1 and G2 using RF algorithm. In part b) of Fig.2, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having the only G1 using the RF algorithm. In part c) of Fig.2, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having without G1 and G2 using the RF algorithm. In part d) of Fig.2, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having only G2 using RF algorithm.



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

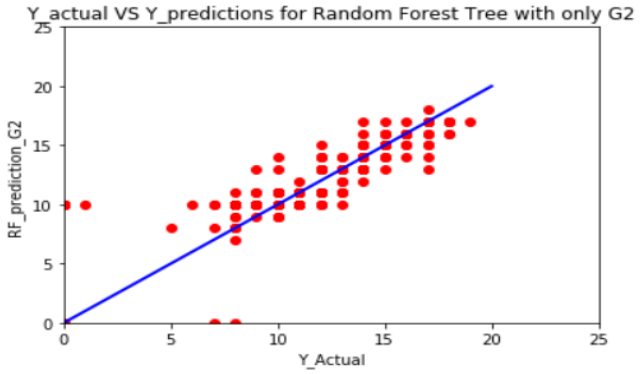


Fig.2. Visualisation of the Actual Value vs Predicted Value using matplotlib library a) Both G1 and G2 known b) Only G1 Known c) Without G1 and G2 d) Only G2 Known by using RF

At last, they calculated the Recall value for all the cases to evaluate performance prediction by applying the classification algorithms in Table III are as follows:

Table III: CASES TAKEN FROM THE ATTRIBUTES FOR THE CALCULATION OF RECALL VALUE

S.No	Classification Algorithms	G1 & G2 Known	Only G1 Known	Only G2 Known	Without G1 and G2	
1.	DT	0.38	0.25	0.35	0.15	
2.	RF	0.42	0.24	0.41	0.16	
3.	KNN	0.32	0.20	0.30	0.12	
4.	SVM	0.43	0.26	0.38	0.16	

From Table III, the authors found the best Recall value for SVM in both the cases when G1 and G2 are known and when only G1 known. But when only G2 known, found that the RF gave the best recall value. They also visualized the prediction between the actual value and predicted value in Fig.3. by using the matplotlib library in python for KNN algorithm. In part a) of Fig.3, there is a visualization between the actual value (AV) on the x-axis and the predicted value (PV) on the y-axis with the case of having both G1 and G2 using KNN algorithm. In part b) of Fig.3, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having only G1 using KNN algorithm. In part c) of Fig.3, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having only G2 using KNN algorithm. In part d) of Fig.3, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having without G1 and G2 using KNN algorithm. In the end, the authors also visualized the prediction between the actual value and predicted value in Fig.4. by using the matplotlib library in python for SVM algorithm. In part a) of Fig.4, there is a visualization between the actual value (AV) on the x-axis and the predicted value (PV) on the y-axis with the case of having both G1 and G2 using SVM algorithm. In part b) of Fig.4, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having only G1 using SVM algorithm. In part c) of Fig.4, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having only G2 using SVM algorithm. In part d) of Fig.4, there is a visualization between the AV on the x-axis and the PV on the y-axis with the case of having without G1 and G2 using SVM algorithm.

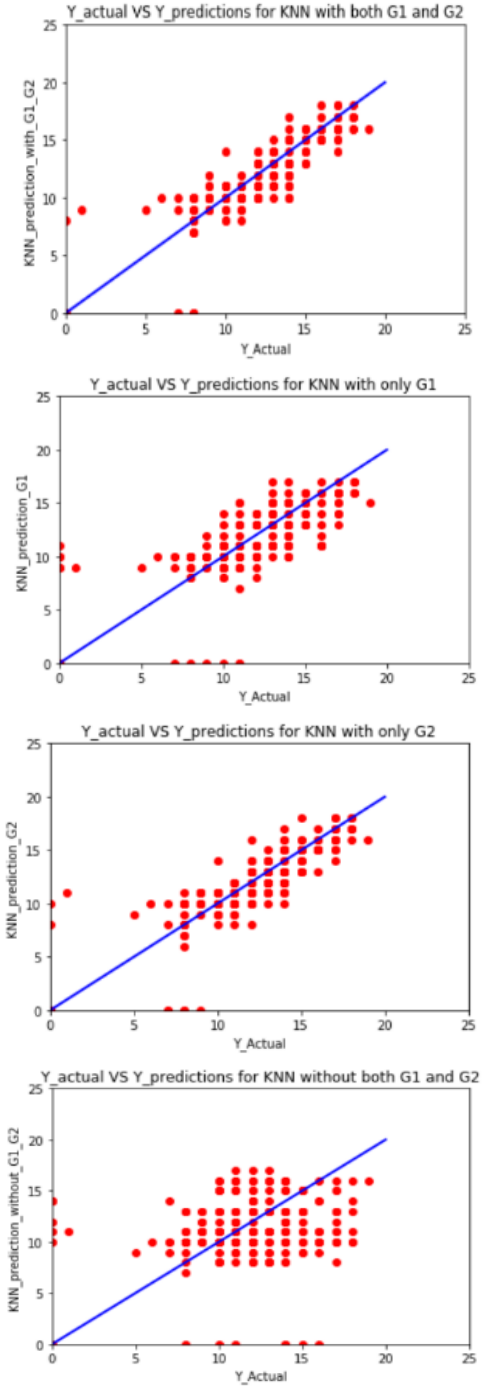


Fig.3. Visualisation of the Actual Value vs Predicted Value using matplotlib library a) Both G1 and G2 known b) Only G1 Known c) Only G2 Known d) Without G1 and G2 by using KNN

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

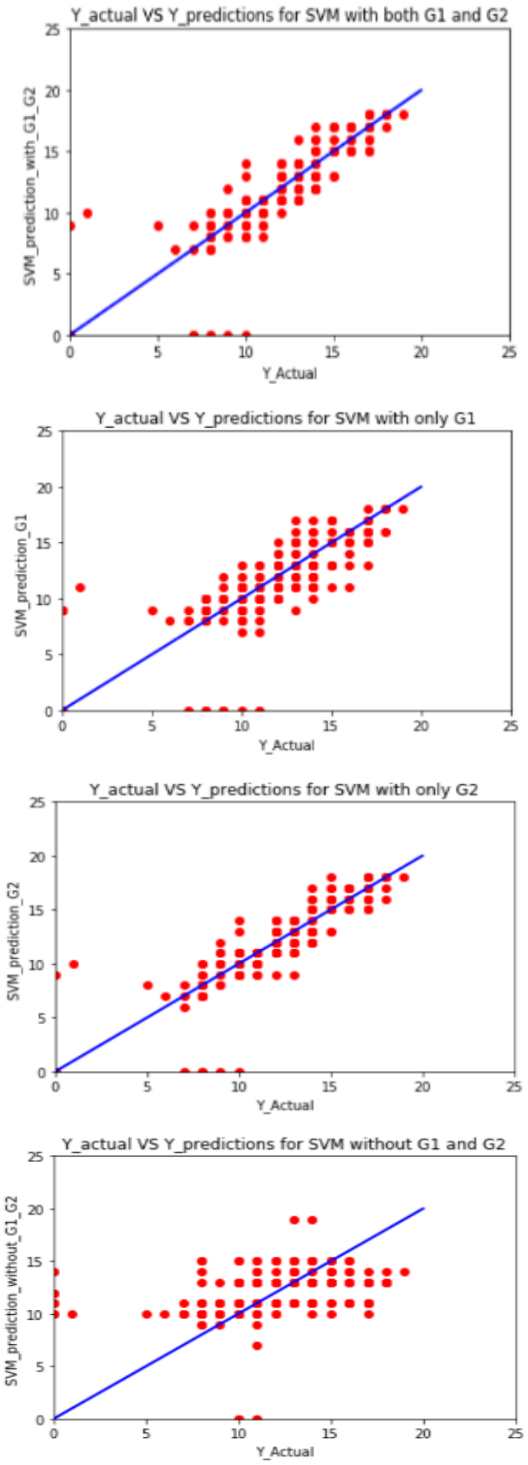


Fig.4. Visualisation of the Actual Value vs Predicted Value using matplotlib library a) Both G1 and G2 known b) Only G1 Known c) Only G2 Known d) Without G1 and G2 by using SVM

V. CONCLUSION

From the experiments the authors conducted, based on the attributes and all the cases they have taken, they can conclude that for POS in terms of accuracy SVM gives the best accuracy when the authors already know the G1 and G2 features but RF performs best in that case even when they do not know about the features G1 and G2. Similarly, when they talk about precision value for POS then SVM again performs better in the known value for G1 and G2 but DT and RF both give the best and same precision value even when the authors do not know about the G1 and G2 features. At the last, the authors have recall value for predicting the POS, in which again SVM performs best when they know the G1 and G2 values but when they do not know the G1 and G2 values then the SVM and RF performs with the equal recall value.

VI. FUTURE SCOPE

For future work, the researchers can do some more experimental results by using regression, association and other ML techniques to get the better accuracy and other evaluation metrics value and the same will be used for predicting the POS in schools, colleges and universities.

REFERENCES

[1] A. Çetinkaya and Ö. K. Baykan, "Prediction of middle school students' programming talent using artificial neural networks," *Eng. Sci. Technol. an Int. J.*, no. xxxx, 2020.

[2] M. Mohammadi, M. Dawodi, W. Tomohisa, and N. Ahmadi, "Comparative study of supervised learning algorithms for student performance prediction," *1st Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2019*, pp. 124–127, 2019.

[3] Rahul and R. Katarya, "A Review: Predicting the Performance of Students Using Machine learning Classification Techniques," *Proc. 3rd Int. Conf. I-SMAC IoT Soc. Mobile, Anal. Cloud, I-SMAC 2019*, pp. 36–41, 2019.

[4] A. Dinesh Kumar, R. Pandi Selvam, and K. Sathesh Kumar, "Review on prediction algorithms in educational data mining," *Int. J. Pure Appl. Math.*, vol. 118, no. Special Issue 8, pp. 531–537, 2018.

[5] B. Sravani and M. M. Bala, "Prediction of Student Performance Using Linear Regression," pp. 1–5, 2020.

[6] A. Rajak, A. K. Shrivastava, and Vidushi, "Applying and comparing machine learning classification algorithms for predicting the results of students," *J. Discret. Math. Sci. Cryptogr.*, vol. 23, no. 2, pp. 419–427, 2020.

[7] E. Tanuar, Y. Heryadi, Lukas, B. S. Abbas, and F. L. Gaol, "Using Machine Learning Techniques to Earlier Predict Student's Performance," *1st 2018 Indones. Assoc. Pattern Recognit. Int. Conf. Ina. 2018 - Proc.*, pp. 85–89, 2019.

[8] A. Jain and S. Solanki, "An Efficient Approach for Multiclass Student Performance Prediction based upon Machine Learning," *Proc. 4th Int. Conf. Commun. Electron. Syst. ICCES 2019*, no. Icces, pp. 1457–1462, 2019.

[9] A. Tripathi, S. Yadav, and R. Rajan, "Naive Bayes Classification Model for the Student Performance Prediction," *2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019*, pp. 1548–1553, 2019.

[10] H. A. Mengash, "Using data mining techniques to predict student performance to support decision making in university admission systems," *IEEE Access*, vol. 8, pp. 55462–55470, 2020.

[11] K. K. and J. S. Raj, "Big Data Analytics for Developing Secure Internet of Everything," *J. ISMAC*, vol. 01, no. 02, pp. 49–56, 2019.

[12] A. Jamil, M. Ahsan, T. Farooq, A. Hussain, and R. Ashraf, "Student performance prediction using algorithms of data mining," *Proc. - 2018 4th Int. Conf. Comput. Eng. Des. ICCED 2018*, pp. 244–249, 2019.

[13] F. J. Kaunang and R. Rotikan, "Students' Academic Performance Prediction using Data Mining," *2018 Third Int. Conf. Informatics Comput.*, pp. 1–5, 2019.

[14] I. Burman and S. Som, "Predicting Students Academic Performance Using Support Vector Machine," *Proc. - 2019 Amity Int. Conf. Artif. Intell. AICAI 2019*, pp. 756–759, 2019.

[15] N. Tomasevic, N. Gvozdenovic, and S. Vranes, "Computers & Education An overview and comparison of supervised data mining techniques for student exam performance prediction," *Comput. Educ.*, vol. 143, no. February 2019, p. 103676, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

[16] R. Hasan, S. Palaniappan, A. R. A. Raziff, S. Mahmood, and K. U. Sarker, "Student Academic Performance Prediction by using Decision Tree Algorithm," *2018 4th Int. Conf. Comput. Inf. Sci. Revolutionising Digit. Landsc. Sustain. Smart Soc. ICCOINS 2018 - Proc.*, pp. 1–5, 2018.

[17] R. Deuja, R. Karna, and R. Kusatha, "Data-Driven Predictive Analysis of Student Performance in College Using Neural Networks," *Proc. 2018 IEEE 3rd Int. Conf. Comput. Commun. Secur. ICCCS 2018*, pp. 77–81, 2018.

[18] V. Nida Uzel, S. Sevgi Turgut, and S. Ayşe Özel, "Prediction of Students' Academic Success Using Data Mining Methods," *Proc. - 2018 Innov. Intell. Syst. Appl. Conf. ASYU 2018*, pp. 1–5, 2018.

[19] Y. S. Alsalman and A. D. Tree, "Using Decision Tree and Artificial Neural Network to Predict Students Academic Performance," *2019 10th Int. Conf. Inf. Commun. Syst.*, pp. 104–109, 2019.

[20] Z. Wu, T. He, C. Mao, and C. Huang, "Exam paper generation based on performance prediction of student group," *Inf. Sci. (Ny).*, vol. 532, pp. 72–90, 2020.

[21] P. Cortez and A. Silva, "Using data mining to predict secondary school student performance," *15th Eur. Concurr. Eng. Conf. 2008, ECEC 2008 - 5th Futur. Bus. Technol. Conf. FUBUTEC 2008*, vol. 2003, no. 2000, pp. 5–12, 2008.

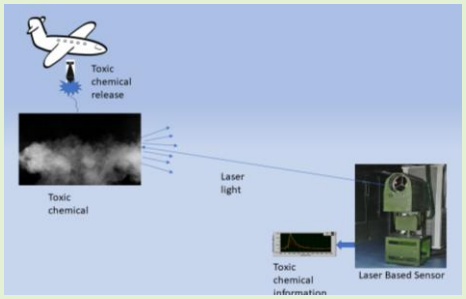
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:55:58 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

Laser Based Systems for Standoff Detection of CWA: A Short Review

M K Jindal, Mainuddin, S Veerabuthiran & A K Razdan

***Abstract*— In recent times, the use of Chemical Warfare (CW) agents against civilian and military by terrorists and rogue countries are recurrent. The examples of such attacks are available in open literature. They are highly dangerous due to their acute toxicity nature. They tend to stimulate and paralyze nerve system of the body and cause other toxic side effects. Once these toxic chemicals are released into the atmosphere, timely detection and identification of the same at standoff distances will help the defensive forces to take proper counter measures. Detection techniques such as ion mobility spectrometry, infrared spectrometry, Raman spectrometry, FTIR, LIDAR etc. have been used widely. Among all these techniques, Raman spectroscopy techniques and LIDAR (LIght Detection and Ranging) is the only active remote detection method that can detect and discriminate the chemical agents at standoff distances. Infrared DIAL (Differential Absorption Lidar) is a versatile technique for remote detection of various chemicals. DIAL technique can provide spatially resolved measurements of these agents with sufficient sensitivity (at few ppm levels) at ranges of several kilometers by exploiting the concept such as scattering and absorption from the atmosphere. The main objective of this review paper is to bring out the details of highly potential chemical warfare agents, their properties and laser based remote detection sensors. LIDAR technique is emphasized in more detail because of their long range detection capability, sensitivity and selectivity. An overview of existing R&D based and commercial systems is also discussed.**



***Index Terms*—Chemical Agents, Chemical Warfare Agents, Differential Absorption LIDAR, Raman Systems**

I. Introduction

THERE HAS BEEN a constant threat of chemical attacks to military personnel as well as to the civilian population. The history of chemical attack is very old. In ancient times poisoned arrows were used as a weapon for the chemical attacks [1]. There have been lots of incidences that were reported where toxic chemicals have been used in the battlefield. Chlorine [2] and phosgene gas [3] has been used from canisters during world war which proved fatal with dispersion and claimed 90000 lives and over one million casualties [4]. Iraq also reported having used chemical weapons containing nerve agents and mustard gas during Iran war in 1980s [5]. During the last three decades, terrorist groups have also started using these chemical warfare agents to create unrest in the civilian population. These trending uses of chemical warfare agents are probably due to the fact that these are easy to manufacture and also cheap to produce. Few examples of chemical attacks are Sarin attack in Tokyo, Japan [6] and bombing in London, UK in 2005 [7]. There have been several other attacks similar to this in the past. Most recent attacks were suspected to be in Syria where chemical weapons have been used on the unprotected civilian populations. Latest have been a chemical attack in Douma city, Syria in April, 2018 which resulted in the death of 43 people [8]. Various chlorinated organic chemicals like dichlorophenol, dichloroacetic acid, bornyl chloride, trichloroacetic acid, chloral hydrate, chlorophenol, etc. have been used in attack [9]. The use of Novichok chemical warfare agents was reported in March, 2018 in Salisbury, UK where Sergey Skripal an ex-Russian spy and his daughter were poisoned [10]

According to Organization for the Prohibition of Chemical Weapons (OPCW), chemical weapon is defined as a weapon that contains any toxic chemical or its precursor that can cause death, injury, temporary incapacitation or sensory irritation through its chemical action. Delivery systems such as aircrafts, Unmanned Aerial Vehicles (UAVs), shells, rockets, etc are used in the targeted area to release chemical agents. The chemical attacks most of the time become very fatal in terms of loss of human lives and long term effects on the exposed population as well as vegetation and animals.

M K Jindal is with Laser Science and Technology Centre (LASTEC), Metcalfe House, Delhi, India -110054 (email: mukeshkumar.jindal@lastec.drdo.in)

Mainuddin is with Deptt. of ECE, JMI, Delhi, India (corresponding author email: mainuddin@jmi.ac.in)

S Veerabuthiran is with Laser Science and Technology Centre (LASTEC), Metcalfe House, Delhi, India -110054 (email: sveerabutriran@lastec.drdo.in)

A K Razdan is with Laser Science and Technology Centre (LASTEC), Metcalfe House, Delhi, India -110054 (email: akrazdan@lastec.drdo.in)

This work is supported by the Ministry of Electronics and Information Technology, Government of India under the Visvesvaraya PhD scheme.

The authors are also grateful to Director, Laser Science and Technology Centre, Metcalfe House, New Delhi, India for his constant support and encouragement for doing research activities in the field of differential absorption LIDAR.

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

Therefore, timely detection and identification of chemical warfare agents at the time of release becomes very important. This is very much required for issuing an early warning for the safety of the population and military. Further, this information will help protection and defense forces to take precautionary actions such as the use of proper antidote, the movement against wind direction, etc so that human loss can be reduced. Systems such as gas chromatography, ion mobility spectrometer, SAW (Surface Acoustic Wave), etc. are used to detect these agents. The drawback of these systems is that one has to introduce these systems in the infected areas. The active standoff systems such as Lidar [11-13] and Raman [14-15] system are preferred because of their remote detection capability of CW agents. These systems can also discriminate the molecules among the group.

Although many authors have written review papers in similar fields, we have focused mainly on various laser based techniques and systems for standoff detection of chemical warfare agents. Kildal et al (1971) [16] had described the laser based methods for remote detection of atmospheric pollutants. But the extended application of these methods for standoff detection of chemical warfare agents is an vital area for homeland security and military establishments. Bogue (2018) [17] had described the techniques and development of systems for remote sensing of chemicals such as CW agents, explosives, air pollutants, industrial leak detection, etc. Bernascolle (2015) [18] had explained briefly about CW agents and standoff detection techniques that include passive and active sensors for detection of chemical warfare agents. He also described commercially available products and compared their performance based on some of the operational parameters. Pacsial-Ong et al (2013) [19] had depicted in detail about the chemical warfare agents, their properties and pharmacological consequences on exposure and also described techniques for point detection and standoff detection of chemical warfare agents along with the details of the systems available. To best of our knowledge, there are no review papers available on laser based sensors for chemical warfare agents detection that describe in detail about techniques and systems.

This review paper presents an overview of laser based techniques and systems for standoff detection of chemical warfare agents. In section II, the classification of chemical warfare agents such as nerve agents, blood agents, choking agents, blister agents, incapacitating agents, etc including their general characteristics and lethal dose are discussed. The detection techniques such as Raman spectroscopy and Differential Absorption LIDAR for standoff measurements of CW agents have been explained in section III. In section IV, Raman Spectroscopy based systems are explained. The most suitable long range remote detection technique is DIfferential Absorption Lidar (DIAL). A review of the systems based on differential absorption lidar has been described in section V. We have extensively covered both commercially available systems and lab based experimentation systems.

II. Classification and Toxicity of Chemical Warfare Agents

Chemical warfare agents are categorized according to their specific use and effect on the human body [20]. Primary forms of chemical warfare agents include nerve agents, blister agents, choking agents, and blood agents.

A. Classification of Chemical Warfare Agents

Nerve agents: Nerve agents are the most dangerous amongst different categories of chemical weapons. They can be either inhaled or absorbed via skin either in liquid or gas form. They inhibit the respiratory and cardiovascular capability of the body as they cause extreme harm to the central nervous system which ultimately can result in death. Tabun, Sarin, Soman, and VX are the most commonly known nerve agents.

Novichok is another organophosphorous chemical warfare agents which is put into nerve agents class[21]. This is a compound group of nerve agents. These agents can go through the personal protective gears and transmit to the body of soldier. Different stages of Novichok synthesis reaction schemes have been given code names as A-230, A-232 and A-234. Novichoks are highly toxic and can be 5-10 order more lethal than V agent VX [22]. The advantage of using Novichok as weapon is that these agents do not degrade easily with passage of time and remain in stabilized state, thus making it more lethal. Recently, in November, 2019 it was added to Chemical Weapon Convention list [23].

Blister agents: They can be present in the form of aerosols, gas or in liquid form and can harm by causing blistering of the skin and severe burns. Inhalation of these agents affects the respiratory system and complications to the digestive tract are caused if ingested. Lewisite, Nitrogen Mustard, Sulfur Mustard, and Phosgene Oximine are the well known blister agents.

Choking agents: Choking agents are chemical toxins, which when breathed in, can attack the respiratory system of the body and can cause respiratory failure. Chlorine, phosgene, and chloropicrin are most common forms of choking agents.

Blood agents: Blood agents interfere with the capability of the body to use and transport oxygen through the bloodstream. Once the blood agents are breathed in, they are absorbed into the bloodstream. Hydrogen Cyanide and Cyanogen Chloride are common forms of blood agents. They are mainly aimed to poison drinks or food. Cyanide salts when mixed with chemicals enhance skin penetration. Hydrogen cyanide (HCN) and cyanogen chloride (Cl-CN) are in liquid form that is transparent-to-pale yellow in color and transforms into gaseous form near room temperature. The exposed victim may feel the burning pain in the eyes. An immediate collapse of any human can happen if high doses of cyanides are used.

Incapacitating agents: These agents when administered in little doses (< 10 mg) can cause the conditions similar to psychotic disorders or other symptoms from the central nervous system (paralysis, rigidity, loss of feeling, etc.) One of the anticholinergic compounds, 3-quinuclidinyl benzilate, was

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

developed as novel chemical agent for battleground use as a psychochemical and assigned the NATO code BZ. The impacts of this agent are transitory and cause incapacitation and helplessness to take decisions.

B. Lethal Dose of CW Agents

Lethal dose, LD_{50}, or LCt_{50} (concentration-time product) is a measure of the lethal dose of a toxic chemical required to kill half of the infected members in given time duration. LD_{50} figures are commonly used as a general gauge of a substance's acute toxicity. A lesser LD_{50} indicates increased toxicity [24]. Lethal Dose and physical properties [20, 25-29] of some of these potential agents are summarized in Table-1.

Chemical warfare agents can be released in different forms such as vapor, aerosol or a combination of both and have different persistence in the atmosphere based on their physical properties. Some CW agents are highly persistent and prove very lethal when released during the chemical attack. The fatality also depends on the characteristics of the delivery system and medium such as atmosphere, water body, soil, etc. After release, the persistence and their lethality effect are governed by their chemical and physical properties, ambient atmospheric parameters like wind speed, temperature, humidity, etc.

Table 1. Physical properties and lethality dose of some chemical warfare agents

Agent		Code	Lethal Dose	Physical Properties
Nerve Agents	Tabun	GA	LCt_{50}: 70 mg-min/m^3 LD_{50}: 21.40 mg/Kg	Colorless to brown liquid Colorless gas Persistence ~ days
	Sarin	GB	LCt_{50}: 35 mg-min/m^3 LD_{50}: 21.40 mg/Kg	Colorless liquid Colorless gas Persistence~ days
	Soman	GD	LCt_{50}:35 mg-min/m^3 LD_{50}: 0.71 mg/Kg	Colorless liquid Colorless gas Persistence ~ day
	VX	VX	LCt_{50}:15 mg-min/m^3 LD_{50}:0.0721 mg/Kg	Amber Oily liquid Persistence ~ weeks to months
	Novichok	A-230 A-232 A-234	Data not available LCt_{50}:6-10 mg-min/m^3 LCt_{50}:15 mg-min/m^3	Normally liquid, volatile, Becomes solid at lower temperature. Does not become solid at lower temperatures and more volatile than A-230. Data not available
Blister Agents	Sulphar mustard	HD	LCt_{50}:900 mg-min/m^3 LD_{50}:100 mg/kg	Colourless to pale yellow liquid
	Nitrogen Mustard	HN	LCt_{50}:1500 mg-min/m^3 LD_{50}: 20 mg/kg	Dark Liquid
	Lewisite	L	LCt_{50}:1400 mg-min/m^3 LD_{50}: 30 mg/kg	Colorless to brown liquid Rapid acting Less persistent than HD
Choking Agents	Phosgene	CG	LCt_{50}:3000 mg-min/m^3	Colorless gas; volatile
	Chlorine	CL	LCt_{50}:6000 mg-min/m^3	Pale yellow-green gas
Blood Agents	Hydrogen	AC	LCt_{50}:5000 mg-min/m^3	Colorless liquid

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

	cyanide		LD_{50}: 100 mg/kg	Highly volatile Rapid detoxification
	Cyanogen chloride	CK	LCt_{50} :11,000 mg-min/m^3	Colorless liquid Highly volatile
Incapacitating Agents	3-Quinuclidinyl benzilate	BZ	LCt_{50}: 2000000 mg-min/m^3	White crystalline solid

III. DETECTION TECHNIQUES

As mentioned earlier, timely detection of CWA is vital to reduce human losses. Several techniques are available for the detection of chemical warfare agents and no single technique can meet all the requirements of detection challenges in terms of range, sensitivity and selectivity. Systems such as SAW, ion mobility spectrometer, gas chromatography, etc. need to be kept physically in the affected area in order to detect the type of chemical agents. This poses the risk of exposing the individual and sensor in the chemical release area. Remote sensing techniques are also useful for advance detection, ranging, identification and concentration estimation of chemical warfare agents. They are classified further as passive and active remote sensing techniques. Passive remote sensing such as Fourier Transform Infra Red (FTIR) receives the light energy from the atmosphere through a wide field of view and further investigates the spectrum of light received through background contrast method in order to detect the chemical cloud in the atmosphere. Active remote sensing sensor uses a radiation source to interact with molecules in the atmosphere. After interacting with molecules, it undergoes some property changes. By analyzing the return signal, one can detect and discriminate the CW agents. In this review paper, only active remote sensing techniques for the detection of chemical warfare agents are considered as these techniques have high resolution and do not depend on the contrast of the background. Hence in this paper two optical active remote sensing techniques, namely Raman spectroscopy and Differential Absorption LIDAR are discussed.

A. RAMAN SPECTROSCOPY

In the Raman spectroscopy technique, the molecular, vibrational, rotational and other low frequency modes are investigated. It is an inelastic scattering based technique in which the wavelength of the transmitted light changes after interaction with the molecule to be probed. The Raman spectroscopy scheme uses a monochromatic laser source that interacts with the chemical species under investigation [30]. The part of the light is scattered in all directions. The incident photons of light are absorbed by the sample and re-radiated at a different frequency than the original incident light. The shift in frequency can be upside or downside which is called Anti-Stokes lines and Stokes lines respectively. By analyzing the shift one can decode the information about vibrational, rotational and other low frequency transitions in the molecule. Hence, Raman is a highly selective fingerprinting technique and used to detect liquid, solid and gaseous samples. Apart from Raman shift, there is Rayleigh scattering whose cross-section is much stronger than the Raman cross-section. Rayleigh scattering is elastic scattering where the reradiated photons are the same frequency as that of the incident photon. Fig 1 shows all the three scenarios in the Raman detection. The light is scattered back at the same frequency as of transmitter due to Rayleigh scattering and can be scattered back at lower frequencies or at higher frequencies which are called Raman Stokes and Anti-Stokes respectively and explained in detail in the subsequent section.

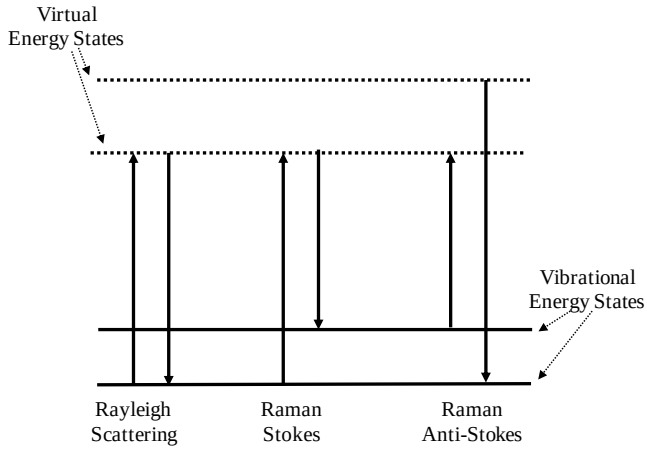


Fig1: Raman energy diagram

Most of the light undergoes Rayleigh scattering while nearly 0.001% light undergoes Raman scattering [31]. Hence the Raman signal is very weak and poses a limitation on standoff detection distance for low energy laser sources. Fig 2 illustrates the basic schematic of Raman sensor setup. The basic setup of Raman spectroscopy sensor consists of a laser transmitter source which is a nanosecond pulsed laser that transmits laser pulses through a beam expander and in the direction of the hazardous sample under investigation. The mostly used wavelengths are 785nm, 532 nm or 355 nm. The scattered radiations consisting of Raman shifted signal, Rayleigh signal and Laser Induced Fluorescence (LIF) signal is collected by a receiving telescope in a coaxial configuration. A lens is used to couple the light from the telescope towards the optical fiber cable. A notch filter is used to cut the incoming Rayleigh scattered radiations. The signal is then given to the entrance slit of the Intensified Charged Coupled Device (ICCD) coupled spectrograph. The stand-off time-

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

gated operation is performed by giving a trigger to ICCD from TTL (Transistor-Transistor Logic) trigger output of the laser source. The gate width of the gating signal is adjusted according to the standoff distance of the sample. The received backscattered Raman signal is processed for identification of sample.Various optical methods like the use of notch filters, laser stop apertures, tunable lasers, etc. are commonly used to reduce Rayleigh scattering and acquire high quality Raman spectra. To enhance the Raman signal, various variants of Raman techniques are used like stimulated Raman spectroscopy, Coherent Anti-Stokes Raman Spectroscopy (CARS), Resonant Raman (RR) and Surface Enhanced Raman Spectroscopy (SERS).

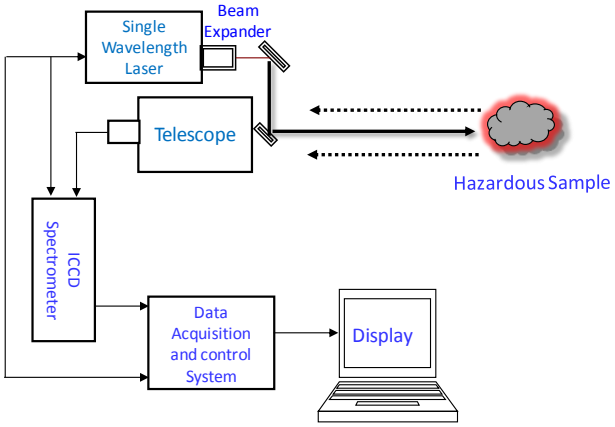


Fig 2: Block diagram of Raman spectroscopy based system

Stimulated Raman Scattering (SRS)

In SRS, very strong laser pulse is used instead of continuous wave laser and a greater portion of incident light gets converted into useful Raman scattering. The difference lies between the spontaneous Raman scattering and Stimulated Raman scattering lies in terms of low divergence and narrow line width of scattered radiation and high intensity [32]. Stimulated Raman technique provides 4-5 orders of signal improvement of Raman signal when contrasted with the spontaneous Raman scattering. The advantages of SRS are coherent strong output and high resolution [33].

Coherent Anti-Stokes Raman Spectroscopy (CARS)

The basic Raman spectroscopic technique has the disadvantage of incoherent scattering and lesser sensitivity. An alternate method is Coherent Anti-Stokes Raman scattering spectroscopy (CARS) which is a nonlinear optical variant of Raman Spectroscopy. In this technique, instead of one laser, two or more collinear strong lasers are used. The frequency of one laser is varied so that the difference in frequency is exactly equal to that of considered Raman active mode. Three frequencies out of which two are at pump frequencies ν_p and one frequency is at Stokes frequency, ν_s is adjusted such that coherent beam frequency (ν_c) is at $\nu_c=2\nu_p-\nu_s$ and $\nu_p-\nu_s$ matches a Raman-active vibrational resonance in the medium. By this, only one strong Raman peak and Raman signal is enhanced with higher intensity [34]. One of the primary requirements in CARS is tunable wavelength laser but with the passage of time many tunable lasers have come out and this technique is being used in many applications [35]. The major advantages of CARS are the increased sensitivity and overcoming the problem of interference from fluorescence as the CARS frequencies are higher than pump frequencies whereas the fluorescence will be at lower frequencies. The major disadvantage of CARS is the higher cost due to instrumentation [36].

Resonance Raman (RR)

In basic Raman spectroscopy technique, when a strong laser beam excites a colored molecule, the light is usually absorbed in that molecule and can create strong fluorescence signal and Raman signal may be masked. To overcome this problem, Resonance Raman which is one of the variants of Raman spectroscopy technique is used. In Resonance Raman Spectroscopy, the percentage of scattered photons is improved by a factor of many orders which can be up to 10^6 times [37]. This is achieved by transmitting laser energy at a frequency which is in resonance with an electronic transition in the atom or molecule. When the resonance phenomena occur, then the vibration and electronic information about the sample under investigation can be obtained. One practical limitation is that the resonance frequencies are different for different substances, thereby a tunable laser needs to be used or many laser sources with fixed wavelengths can be used for the detection of various agents from standoff distances [38]. The advantage of Resonance Raman technique is that it is a much more sensitive and selective technique than the basic Raman technique because only Raman bands of the chromophore in tune with the excitation wavelength are enhanced.

Surface Enhanced Raman Spectroscopy (SERS)

The weak cross-sections problem in basic Raman that very small percentage of light being Raman shifted and rest of light is elastically scattered can be resolved by using another variant of Raman spectroscopy which is called Surface-Enhanced Raman Spectroscopy (SERS). In Surface-Enhanced Raman Spectroscopy, the resonance effect is created by surface-induced resonance(s) instead of electronic transitions intrinsic to the molecule [39]. In Resonant Raman, the electronic transition between molecule and metal surface is electrically connected and is called chemical enhancement effect. The molecules are absorbed on certain metal surfaces and Raman signal is 5-6 orders of magnitude greater than that from the same molecule in bulk volume. Mostly the surfaces used for this purpose are electrochemically etched silver electrodes as well as silver and gold colloids. The advantages like such as fast response, ultra sensitivity multiplexing ability and comparatively low cost are making this technique popular for a number of applications. Also, SERS has the capability of detection of chemical warfare agents due to the combination of high label-free sensitivity and molecular fingerprinting ability [40].

The advantage of Raman technique is that it uses only a single wavelength laser. Another advantage is the detection of suspect molecules in the presence of interfering molecules

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

since the Raman lines of a molecule are specific to them. The major disadvantage is sensitivity as the standoff distance increases. For increasing standoff distance, one has to increase the laser power which can pose many hazards like eye safety, personnel safety, etc. and also the size of the system may increase with the power of the laser.

B. D IFFERENTIAL ABSORPTION LIDAR (DIAL)

Differential Absorption LIDAR (DIAL) is one of the LIDAR techniques widely used for standoff detection of CW agents, pollutants and Toxic Industrial Chemicals (TICs) in the atmosphere [41-46]. This is an elastic scattering and absorption based technique wherein changes in properties of the signal determine the agent present in the probed area and the concentration of the selected agent. This technique involves high energy tunable pulsed laser system. Two pulses at different wavelengths that are closeby in wavelengths are sent into the atmosphere. One wavelength is called online wavelength which is highly absorbed by the agent under probe and another wavelength is called offline wavelength is not absorbed or absorbed relatively less by agent under probe as compared to that of online wavelength. As a result, the intensity of the backscattered return signal at online wavelength will be less as compared to return signal intensity at offline wavelength. The ratio of intensities along with other parameters is used to calculate the concentration of probed agent. The time of flight of the return signal in reference to the start of transmittance of laser pulse into the atmosphere is used to determine the distance of agent. The signal at various range bins is used to calculate the concentration at different distances to estimate range resolved concentration profile. Fig 3 shows the basic scheme for DIAL measurement setup. The setup consists of a tunable pulsed laser as Lidar transmitter source which transmits laser pulses in the region of interrogation of the atmosphere at fixed pulse repetition frequency. The laser can be tuned to online and offline wavelength alternatively. The laser beam is expanded using beam expander to reduce beam divergence so as to cover the probe area of interest. Before transmitting the expanded beam to the atmosphere, it is folded twice in order to make transmitter and receiver coaxial. The backscattered radiations from the toxic cloud are collected in the field of view (FOV)

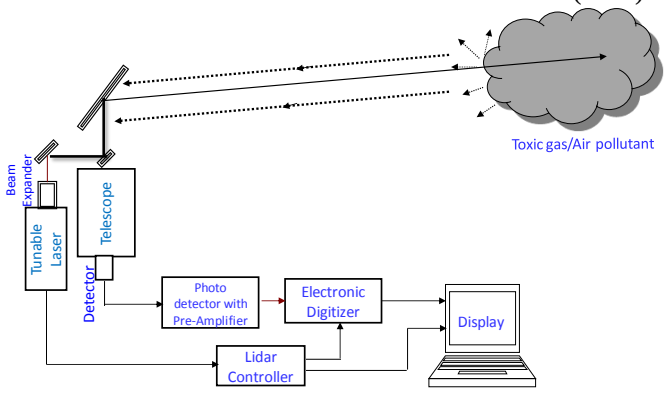


Fig 3: Block Diagram of Differential Absorption LIDAR System

of the receiver telescope which ultimately focuses the received radiations to an optical detector. The choice of the optical detector depends on the Transmitter wavelength band, detectivity, gain, dark current, etc. The electrical signal from the detector is amplified using a preamplifier and converted in digital signal using a data acquisition system. The digitized signal is processed to estimate the distance and concentration of chemical warfare agent. This information is finally displayed on a display unit. This technique is highly sensitive and selective. Chemicals absorb wavelengths selectively and by choosing the appropriate wavelengths one can discriminate the agents easily. The selection of appropriate online and offline wavelengths is very important and is done by taking into account various factors like absorption cross-sections at online and offline wavelengths, vicinity of online & offline wavelengths and presence of atmospheric interfering molecules.

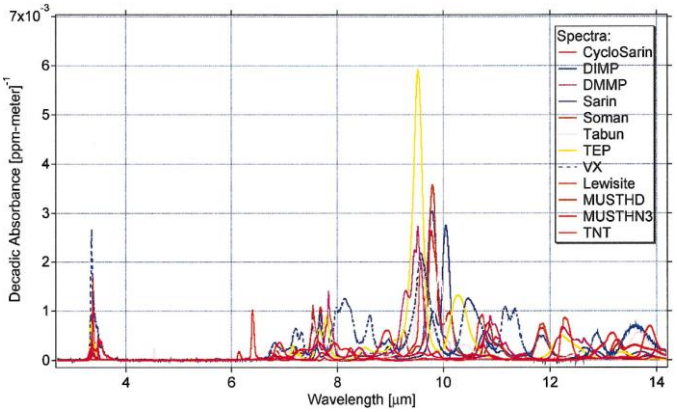


Fig 4**:** IR Absorption Spectra of CW agents "Reproduced from [47], with the permission of AIP Publishing"

Spectral absorption characteristics of chemical warfare agents is an essential factor that has to be considered while designing a DIAL for remote detection of CWA as the maximum achievable range, sensitivity and selectivity is also governed by this information. Fig 4 [47] shows typical absorption spectra of potential chemical warfare agents. As seen from the figure, the majority of CW agents (nerve agents, blister agents, toxic industrial chemicals, etc.) have absorption of light in a wavelength band of 3μm- 14 μm [47]. It can be further seen that within 3μm-14μm band, these agents have two distinct absorption bands i.e. 3μm- 4 μm and 9μm- 11 μm wavelength region [48-49]. These wavelength bands with requisite high laser pulse energies cannot be present in one laser source. 9μm-11μm wavelength band is catered by CO_2 lasers and has been used for standoff detection of chemical warfare agents [13,50-52]. 3μm-4μm wavelength band is catered by OPO (Optical Parametric Oscillator) based lasers which use solid state lasers as their pump source and have been used for standoff detection of chemical agents [53-54]. It is clearly understood that the differential absorption Lidar systems operating in the above spectral band have the capability of detecting these agents at standoff distance.

Similar to DIAL technique, two wavelengths (266 & 355

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

nm) based laser induced fluorescence (LIF) technique have also been reported for detection of chemicals such as fuel, pesticides, etc. Kraus et al (2019) [55] reported the online discrimination of chemical substances using this technique. The authors have described predictive models and algorithms for identification of chemicals within a substance.

IV. SYSTEMS BASED ON RAMAN SPECTROSCOPY TECHNIQUE

There are many systems available in the market based on Raman Spectroscopy technique. These systems along with their functional features are explained in the subsequent section. Few R& D systems developed by research groups are also explained.

A. RespondeR RCI by Smiths Detection

RespondeR RCI is a Raman spectroscopy based sensor that is portable and can be taken to field and commercially available by M/s Smiths Detection, USA [56]. It uses a diode laser as a transmitter with 75 mW output laser power at 785 nm wavelength which is focused on the sample under investigation. The scattered light is received by an optical receiver and a holographic diffraction grating based spectrometer is used in wave-number band from 225 cm^{-1} to 2400 cm^{-1} with a resolution of 12 cm^{-1} and a CCD spectrometer is used to measure the intensity of the spectrum. The spectrum from the sample is correlated with the spectrum in the library and the correlation score is calculated and hence chemical is identified. The system is light weight (weighs around 6 pounds), compact and detection time mentioned is less than 30 secs [57]. It serves the purpose of detection and identification of chemicals, explosives, TICs, etc [58].

B. Thermo Scientific FirstDefender RM

Thermo Scientific FirstDefender RM is manufactured by Thermo Scientific, USA (formerly known as Ahura Scientific). FirstDefender is a lightweight, handheld and rugged spectrometer based on Raman Scattering [59]. In this system, a laser light is sent to the suspected target. Due to vibrational bonds of the molecules, light undergoes absorption, reflection and scattering. The backscattered shifted wavelengths are taken and presented as a Raman spectrum. This spectrum is correlated against a chemical database (library) and the chemical is recognized. A laser having variable power options of 75 mW, 125 mW and 250 mW is utilized for a transmitter source at 785 nm. A dispersive spectrometer and charge coupled detector is used for generating wavelength spectrum with a spectral range of 250 cm^{-1} to 1875cm^{-1} with 7 cm^{-1} to 10.5 cm^{-1} resolution. Its weight is 1.8 pounds and the sample exposure time can be set as manual or automatic with a minimum exposure time of 5 ms [60]. It has the capability to detect explosives, chemical warfare agents (CWAs), toxic industrial chemicals (TICs), etc. [61]. The main limitations properly understand the chemical type and optimizing the detection distance of the sample so that laser beam can be focused properly [62].

C. Raman Sensor by LASTEC

Gulati et al (2017 & 2019) [63-64] from Laser Science & Technology Centre, Delhi, India demonstrated the experimental system based on time gated Raman spectroscopy for the detection of several explosives from standoff distances of 10m. The setup was built using a 532 nm, 4 ns pulse duration laser source which emits the pulses at 10 Hz. The Raman spectral signal is received by an 8-inch telescope along with a long pass notch filter and recorded by a gated ICCD spectrometer which is thermoelectrically cooled. The system is capable to acquire weak Raman spectra suppressed in high background noise. The system can be used for in-situ real time detection of hazardous chemicals also.

Gupta et al (2015) has used multiple beams for Raman detection of samples of acetonitrile and urea. A diode pumped solid state CW laser with power of 450 mW and wavelength 532 nm was used as excitation source. Raman signal was collected from a distance of half a metre by one inch receiving optics having focal length of 50mm. The received radiations were focused on a spectrometer after passing through a notch filter at 532 nm [65].

V. SYSTEMS BASED ON DIFFERENTIAL ABSORPTION LIDAR TECHNIQUE

Differential Absorption Technique for the detection of CW agents is a very important technique and used by many groups worldwide. Some groups have made prototype experimental systems and several commercialized systems are also available. The following are the details of some of the systems developed by groups/firms internationally.

A. FALCON-4G

Falcon 4G system (as shown in Fig 5) is available commercially for standoff detection of CW agents and the product is manufactured and sold by SEC Technologies, Slovakia [66]. The system is working on Differential Absorption LIDAR technique. Two tunable pulsed CO_2 lasers sources operating in 9.2 to 10.7 µm wavelength band are being used as a laser radiation source. The lasers are pulsed and can

Fig 5: FALCON-4G LIDAR system

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

be tuned to 70 different wavelengths to have wide coverage of CW agents and TICs. The pulsed energy used is 50 mJ (at 10P20) with a pulse repetition rate of less than 4 Hz. The instantaneous field of view of the receiver is 1 mrad. A pan and tilt platform is used for facilitating the permanent surveillance of the area of interest. The system is capable of detecting and indentifying the CWAs and TICs. from a standoff distance upto 6 km. It can be mounted on a vehicle with automatic 360° scanning capability, or used from ground on the tripod. The system can detect chemical warfare agents such as GA, GB, GD, GF, VX, HD and L. The reported detection threshold concentration is as low as 0.05 mg/m^3 for GA agents, 0.10 mg/m^3 for GB, 0.15 mg/m^3 for GD, 0.40 mg/m^3 for GF, 0.10 mg/m^3 for VX, 1.25 mg/m^3 for HD. The system is also capable of detecting various TICs (Toxic Industrial Chemicals) like acetic acid, ammonia, diethylether, toluene, ethanol, ethylene, formic acid, gasoline, methanol, SF_6, Triethyl Phosphate (TEP), 2-(N-Morphonilo)-ethanesulfonic acid (MES) and others. The system has expandable database. The system takes 0.25 secs for the detection of single agent in particular along the line of sight. The system can give concentration profile upto 1.5 km and have a detection range of 6 km. The system has been reported to be tested successfully by TNO with the detection of less than 150 grams of SF_6 from a standoff distance of 5 kms [67].

B. DIAL System Developed by Italy

The system (Fig 6) developed by the Italian group under the project BONAS is based on Differential Absorption Lidar technique. The system contains a laser transmitter which is a combination of OPO and OPA and gives two wavelengths namely signal and idler. Idler wavelength is used as transmitter wavelength for investigating the atmosphere [68-71]. The laser is wavelength tunable in the span of 3.0 µm to

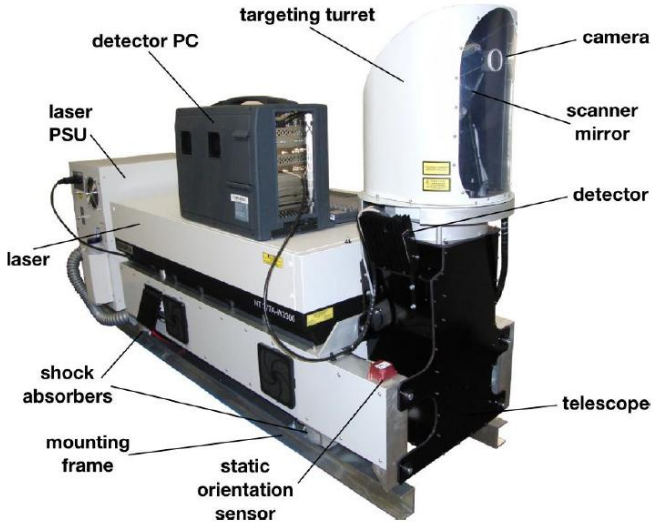


Fig 6: BONAS DIAL System

3.5 µm with linewidth of 5 cm^{-1} and emits an elliptical beam having a divergence of 1.1 mradx1.25 mrad, and beam size of 22 mm. The transmitter emits laser pulses at a pulse repetition rate of 10 Hz with the energy at 12 mJ per pulse. There is a scanning mirror which is common to transreceiver chain and is part of telescope and the detector (Vigo, Poland PVI-4TE) having a transimpedance amplifier gain of 10^4 V/A and bandwidth of DC-50MHz. A scanner is mounted on top of the laser transmitter and receiver which are coaxially aligned. The scanner is capable of covering almost 360 degrees horizontally and ±30 degrees vertically. The Data acquisition system is a 14 bit, 125 MS/sec digitizer board with 2 channels from GAGE. The system is a transportable system and is mounted in a van. This system reported is an experimental system and have been used for detection of acetone.

C. Mobile Lidar Complex by Laser Systems, Russia

Mobile Lidar complex developed by Laser Systems, Russia is a mobile platform based LIDAR system and has many lidar channels operating in the wavelength band of 266 nm to 11 µm. It has an aerosol Lidar channel, IR DIAL channel and UV DIAL channel. IR Differential Absorption Lidar channel is based on the gas CO_2 laser which has 60 different laser lines and transmits laser pulses into the atmosphere. The laser has the maximum energy of 500 mJ per pulse at 10P20 line with a pulse width of 1-2 microseconds and pulse repetition frequency of 10 Hz. In this system, two coherent channels are operated in succession in a time interval of 10 msec with a repetition rate of 10 Hz. The backscattered radiations are collected by a receiver telescope that has a Meresenne configuration with an aspheric primary mirror of an optical diameter of 500 mm. The detector used is MCT detector which gives an analog signal to electronics consisting of an analog to digital converter of 100 MHz bandwidth. The system has a physical scanning capability with the help of a motorized scanner system mounted on the top of the vehicle and can cover a scanning angle of -7 degrees to +15 degrees in elevation and ±180 degrees in azimuth with a minimum scanning step of 40". The system is operated by automated control software. It has the GPS system to map the concentration according to geographical coordinates [72]. The maximum detection range for the IR DIAL channel is 15 km.

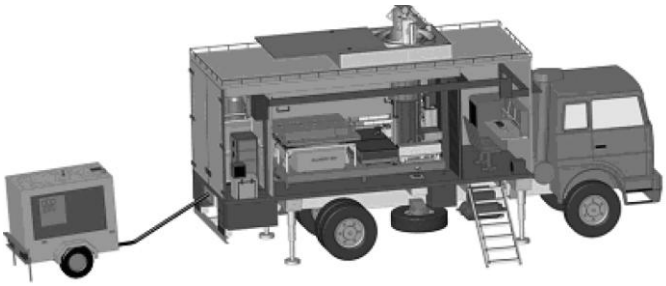


Fig 7: Layout of Mobile LIDAR Complex

The system has the functionality of providing information on the concentration and physical nature of aerosols in the atmosphere, measurement of concentrations of chemicals used in military attack, etc. Fig 7 shows the layout of Mobile Lidar Complex [73]. A similar system is also sold commercially by Laser Systems, Russia [74], and claim to have a minimum detection range of 0.5 km and a maximum range of 5 km and is capable of detection of hydrocarbons, Freon gases,

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

organophosphates (chemical weapon). The scanning range is from 15 to +90 degrees in vertical and +/- 180 degrees in horizontal with a maximum scanning speed of 20 degrees per second with a resolution of 40 arc seconds. The system is mounted on a truck for mobility and field operations.

D. OPO DIAL System by V.E. Zuev Institute of Atmospheric Optics, Russia

V.E. Zuev Institute of Atmospheric Optics SB RAS, Russia has designed an experimental Differential Absorption LIDAR system based on Optical Parametric Oscillator pulsed laser. The OPO Laser (SOLAR LS, Minsk, Russia) can transmit wavelengths in spectral bands from 1.4 µm 1.6 µm, 1.8 µm to 2.6 µm and 3.0 µm to 4.0 µm with pulse energies of 6 mJ, beam size 4 mm at a repetition rate of 10 Hz. The pulse width is 10 to 13 nsec and spectral line width less than1 cm^{-1}. The transmitter also consists of a beam expander which reduces the beam divergence to 100 µrad and the beam size becomes 40mm. The receiver consists of a 300 mm diameter primary mirror cassegrain telescope with 0.0017^{o} field of view and a thermally cooled IBSG photodiode (IBSG, St.Petersburg, Russia) having detectivity, (3–6)x10^{9} cm.$Hz^{1/2}W^{-1}$ and active diameter of 0.3 mm for the transformation of the optical signal into an electrical signal. The data acquisition is provided by Rigol 1104Z-S and Acute TS2202E oscilloscopes with a sampling speed of 1 GS/sec. The system has been used to study the distribution of methane concentration in the atmosphere [75-76]. Romanovskii et al (2018) [77] and (2019) [78] described the details of Mid IR wavelength OPO laser based Differential absorption Lidar and experiments were conducted for the detection of atmospheric CO_2, H_2O and CH_4. Although the system has been reported only for the detection of various atmospheric gases, the system has all the instrumentation with requisite specifications that can be used for the detection of a lot of potential chemical agents that fall in 3.0 µm to 4.0 µm wavelength region.

E. Differential Absorption LIDAR system by NPL, India

An experimental Differential Absorption LIDAR system has been by National Physical Laboratory, India and used for measurement of minor constituents of atmosphere like water vapor, ethylene, ammonia and ozone[79]. This system utilizes CO_2 laser in the waveband of 9-11 µm (Edinburgh Instruments) with a power of 2-3 Watt and beam divergence of 10±1 mrad. The system uses a spectrum analyzer from optical engineering to validate the transmitted wavelengths.

F. DIAL System by University of Rome Tor Vergata, Italy

University of Rome Tor Vergata, Italy has developed a compact multi wavelength Differential Absorption LIDAR system in far infrared band using mini TEA CO_2 pulsed laser [80-81]. The laser (TEA MIL-5) is a tunable laser capable to tuning up to 66 different wavelength lines in the wavelength band of 9-11µm through a stepper motor integrated with grating. The laser is able to fire pulses a pulse repetition frequency of 100 Hz and pulse width of 50 nsec and per pulse maximum energy of 50 mJ. The maximum energy of laser is available at 10P60 line. The laser pulses are fired into the open atmosphere and backscattered laser radiations are collected by a Newtonian type optical receiver telescope having primary mirror diameter of 200 mm and subsequently focused into a liquid nitrogen cooled Mercury Cadmium Tellulide (MCT) detector. The analog signal from detection system is digitized using 14-bit resolution, 100 MSPS digitizer card, stored in a PC and transmitted wirelessly to a remote workstation. The system is mounted on a motorised equatorial mount to direct the laser beam in the area of measurement. LabVIEW based software has been used for data acquisition and analysis purpose. The system incorporates a multi-wavelength scheme where more number on set of on and off wavelengths for DIAL detection of chemicals, thereby improving the concentration calculation accuracy. The system has the capability of detection and identification of 24 chemicals. The system has been reported to be used for detection of water vapors and ammonia and water vapors from standoff distance of 100 m [81] and 500m [82].

G. CO_2 Laser Based Differential Absorption Lidar by LASTEC, India

A compact trolley-mounted CO_2 DIAL System has been realized at Laser Science and Technology Centre (LASTEC), Delhi, India as shown in Fig 8. The system consists of a pulsed TEA CO_2 Laser (M/s OptoSystems Ltd., Russia) which is tunable and emits ~50 mJ energy (highest energy at peak wavelength). The laser has the tunability in the wavelength region of 9.2–10.8 μm and can emit pulses of 2 μs pulse width at a programmable PRF from 10 to 100 Hz. The DIAL receiver system consists of a 200 mm beam reducing telescope

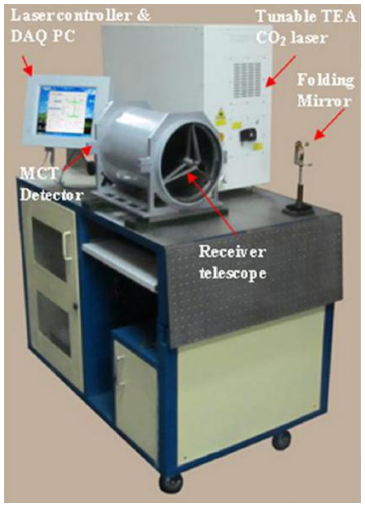


Fig 8: Experimental CO_2 DIAL System developed by LASTEC, Delhi, India

together with a focusing lens of 25.4 mm diameter in size, focal length of 15 mm and a liquid nitrogen (LN_2) cooled MCT detector with an active area 0.1 mm× 0.1 mm, and a 9–11 μm wavelength band optical filter. The Data Acquisition system is a National Instruments, 12-bit voltage resolution, 10 MS/s sampling DAQ card. The system has the capability of stand-off detection of chemical clouds in aerosols as well as in vapor form upto around 200 m range in the atmosphere. The system was reported to be tested for the detection of diethyl

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

ether (DEE) (a Toxic Industrial Chemical) from remote distances upto ~100 m [83].

H. OPO DIAL System developed by LASTEC, India

LASTEC has also developed an OPO DIAL system for the detection of chemicals at standoff distances. Fig 9 shows the photograph of the system. The system uses a tunable OPO laser source operating in 3.0 to 3.45 µm wavelength band as transmitter source with the wavelength tunability at an interval of 1nm. The output laser pulses have the maximum pulse energies of 5 mJ per pulse, with 6 ns of pulse duration and a pulse repetition rate of 10 Hz. A 200 mm size (diameter) Cassegrain telescope focuses the collected radiation onto an MCT detector. It has visible alignment laser to align with the infrared laser beam. The system has the capability to perform physical scanning with scanning speeds upto 10 degrees per second and can perform scanning in azimuth and elevation with limits between -60 degrees angle to +60 degrees angle in azimuth direction and -10 to 30 degrees angle in elevation. The system has also collocated CCD to assist the operator to see the line of sight scene. The system is a mobile system and supported by the shock absorbers to movement for field testing. The system normally runs on AC, 230V power but has a captive power source of 5kVA online Uninterrupted Power Supply (UPS) for power backup.

Fig 9: OPO DIAL System by LASTEC, India

The system has been reported for the detection of Thiodiglycol (TDG) smoke cloud, which is a simulant of chemical warfare agents at a distance of 1 km with a concentration profile at different dispersion stages [11]. The system has also been reported for the detection of ambient atmospheric methane [84].

VI. CONCLUSION

It is clearly understood that CW agents can be used against civilian as well as military establishments by terrorists and rogue countries. Considering the acute toxicity of the CWA, it is strongly emphasized the timely and prompt detection and identification of CWA becomes the need of the hour. Standoff detection sensors are superior compared to other sensors because of their strength of detecting and indentifying the chemical agents at far off distances. In this paper, we have discussed a brief of existing laser based sensors such as Raman sensors and DIAL sensors for the detection of CW agents. Both techniques prove to have the capability of detecting and discriminating chemical warfare agents. Laser based sensors are developing vigorously because of advanced research development in the area of laser, detectors, control electronics, etc. The important criteria such as portability, response time, detection and discrimination capability, etc are the most desirable parameters of standoff chemical sensors. Laser based sensors that are currently available in the market and under development have different advantages as well as disadvantages. No single sensor can meet all the desired parameters. A sensor that is good for one type of CWA may not be suitable for another type depending upon the operational scene and nature of CWA present in the scene. For example, Raman sensor cannot detect chemical vapors and aerosols at longer distances. Also, it is incapable of detecting CWA present in nontransparent medium like munitions. It is also reported that it produces false positive and false negative events and failed to detect low concentration in the contaminated area. DIAL method uses multiple wavelengths for the detection of various agents and has range capability upto few kilometers. The limitations of presently available Lidar sensors are that they work well in static mode rather than mobile for detection mode of CWA. Also, the performance is affected by weather conditions such as dust, rain and snow. Highest sensitivity and selectivity of the sensors can be achieved easily due to improvement in recent advancements taking place in these technologies. Over the past few years, significant progress in research and development is taking place globally to improve the detection capability of laser based systems even in the contaminated area. We believe that numerous lidar products with improved performance will be available in the commercial market in near future.

REFERENCES

[1] A. Mayor, “Greek fire, poison arrows, and scorpion bombs: biological & chemical warfare in the ancient world”, Penguin, 2008.

[2] W. van Der. Kloot, "April 1915: Five future nobel prize–winners inaugurate weapons of mass destruction and the academic–industrial–military complex”, Notes *and Records of the Royal Society of London* 58, no. 2, pp. 149-160, 2004.

[3] C. A. Gibson, "Chemical Warfare as Developed During the World War—Probable Future Development" ,*Bulletin of the New York Academy of Medicine* , vol. 13, no. 7, p. 397, 1937.

[4] O. K. H. Spanjaard and O. Khabib, "Chemical weapons." Terrorism and public health: a balanced approach to strengthening systems and protecting people”, p. 199, 2003.

[5] J. Ali, "Chemical weapons and the Iran-Iraq war: A case study in noncompliance", *The Nonproliferation Review*, vol. 8, no. 1, pp. 43-58, 2001.

[6] T. Okumura, N. Takasu, S. Ishimatsu, S. Miyanoki, A. Mitsuhashi, K. Kumada, K. Tanaka, and S. Hinohara, "Report on 640 victims of the Tokyo subway sarin attack", *Annals of emergency medicine* 28, no. 2, pp. 129-135, 1996.

[7] J. H. Banoub and R. M. Caprioli, eds,” Molecular *Technologies for Detection of Chemical and Biological Agents”*, Springer, 2017.

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

[8] G. Koblentz, "Chemical-weapon use in Syria: atrocities, attribution, and accountability", *The Nonproliferation Review*, vol. 26, no. 5-6, pp. 575-598, 2019.

[9] OPCW Technical Secretariat, "Report of the fact-finding mission regarding the incident of alleged use of toxic chemicals as a weapon in douma, syrian arab republic, on 7 april 2018", *Organisation for the Prohibition of Chemical Weapons*, S/1731/2019, march 2019.

[10] K. Kuca and E. Nepovimova, "Are we facing NOVICHOK nerve agent threat?", *Australasian Medical Journal (Online)*, 12(2), pp. 49-52, 2019.

[11] S. Veerabuthiran, A. Razdan, M. Jindal and G. Prasad, "Open field testing of mid IR DIAL for remote detection of thiodiglycol vapor plumes in the topographic target configuration", *Sensors and Actuators B: Chemical*, vol. 298, p. 126833, 2019.

[12] M. Tehrani, M. Mohammad, E. Jaafari and A. Mobashery, "Setting up a mobile Lidar (DIAL) system for detecting chemical warfare agents", *Laser Physics*, vol. 25, no. 3, p. 035701, 2015.

[13] C. B. Carlisle, J. E. Van der Laan, L. W. Carr, P. Adam and J. P. Chiaroni, "CO_2 laser-based differential absorption lidar system for range resolved and long range detection of chemical vapor plumes" *Appled Optics*, vol 34, pp. 6187-201, 1995.

[14] F. Yan, M. Wabuyele, G. Griffin, A. Vass and Tuan Vo-Dinh, "Surface-enhanced Raman scattering detection of chemical and biological agent simulants", *IEEE Sensors Journal*, vol. 5, no. 4, pp. 665-670, 2005.

[15] S. D. Christesen, J. M. Lochner, A. M. Hyre, D. K. Emge, and J P Jones. "UV Raman spectra and cross sections of chemical agents.", *Chemical and Biological Sensing VII*, International Society for Optics and Photonics, vol. 6218, p. 621809, 2006.

[16] H. Kildal and R. Byer, "Comparison of laser methods for the remote detection of atmospheric pollutants", *Proceedings of the IEEE*, vol. 59, no. 12, pp. 1644-1663, 1971.

[17] R. Bogue, "Remote chemical sensing: a review of techniques and recent developments", *Sensor Review*, vol. 38, no. 4, pp. 453-457, 2018.

[18] P. F. Bernascolle, "CWA stand-off detection, a new figure-of-merit: the field surface scanning rate", in *Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XIV*, vol. 8710, p. 871008, International Society for Optics and Photonics, 2013.

[19] E. Pacsial-Ong, "Chemical warfare agent detection a review of current trends and future perspective", *Frontiers in Bioscience*, vol. 5, no. 2, pp. 516-543, 2013.

[20] M. Kloske and Z. Witkiewicz, "Novichoks – The A group of organophosphorus chemical warfare agents", *Chemosphere*, vol. 221, pp. 672-682, 2019.

[21] P. Chai, B. Hayes, T. Erickson and E. Boyer, "Novichok agents: a historical, current, and toxicological perspective", *Toxicology Communications*, vol. 2, no. 1, pp. 45-48, 2018.

[22] S. Costanzi and G. D. Koblentz, "Updating the CWC: How We Got Here and What Is Next", *Arms Control Today*, vol 50, No 3, pp 16-20, 2020.

[23] F. R. Sidell, "Chemical warfare agents." *Emergency Toxicology",* Philadelphia, USA: Lipincott-Raven Publishers, 1998, pp. 255-266.

[24] T. T. Marrs, R. L. Maynard, and F. Sidell, eds. "*Chemical warfare agents: toxicology and treatment",* John Wiley & Sons, 2007.

[25] "Brief Description of Chemical Weapons, Organisation for the Prohibition of Chemical Weapons", *Organization for the Prohibition of Chemical Weapons*, October 21, 2014.

[26] J. E. Forman, W. Alwan, W, F. Nurul, I. S. Sun and V. Suri, "A visual guide to science & technology relevant to the chemical weapon convention", Organisation for the Prohibition of Chemical Weapons, OPCW Reprographic Centre, 2018

[27] US Army Medical Research Institute of Chemical Defense. Chemical Casualty Care Division. *Medical Management of Chemical Casualties: Handbook*. Chemical Casualty Care Division, US Army Medical Research Institute of Chemical Defense, 1999.

[28] L. Wright, C. Pope, and J. Liu, "The Nervous System as a Target for Chemical Warfare Agents", in *Handbook of Toxicology of Chemical Warfare Agents*, pp. 463-480. Academic Press, 2009.

[29] E. Nepovimova and K. Kuca, "Chemical warfare agent NOVICHOK - mini-review of available data", *Food and Chemical Toxicology*, vol. 121, pp. 343-350, 2018.

[30] G. Singhal, Mainuddin, R. Rajesh, M. Beg, R. Tyagi and A. Dawar, "Overview of Optical Techniques for Characterization of High-Power Infrared Gas Lasers", *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4165-4173, 2015. Available: 10.1109/jsen.2015.2409114.

[31] Raman Spectroscopy Basics by Princeton Instruments, [Online]. Available: https://pdf4pro.com/cdn/raman-spectroscopy-basics-portland-state-university-5048f3.pdf, [Accessed: 19- Apr- 2020].

[32] M. Maier, "Applications of stimulated Raman scattering", *Applied Physics*, vol. 11, no. 3, pp. 209-231, 1976.

[33] A. Owyoung, "Sensitivity limitations for CW stimulated Raman spectroscopy", *Optics Communications*, vol. 22, no. 3, pp. 323-328, 1977.

[34] A. Portnov, S. Rosenwaks and I. Bar, "Detection of particles of explosives via backward coherent anti-Stokes Raman spectroscopy", *Applied Physics Letters*, vol. 93, no. 4, p. 041115, 2008.

[35] P. Chen, C. Joyner, S. Patrick and K. Benton, "High-Speed High-Resolution Gas-Phase Raman Spectroscopy", *Analytical Chemistry*, vol. 74, no. 7, pp. 1618-1623, 2002.

[36] J. R. Ferraro, *Introductory raman spectroscopy*. Elsevier, 2003.

[37] E. Smith and G. Dent, *Modern Raman spectroscopy: a practical approach*, John Wiley & Sons, 2019.

[38] G. Rustad, "Technologies for standoff detection of liquid chemical warfare agents", *Norwegian Defence Research Establishment*, FFI-rapport 2011/01890, 2011.

[39] A. Hakonen, P. Andersson, M. Stenbæk Schmidt, T. Rindzevicius and M. Käll, "Explosive and chemical threat detection by surface-enhanced Raman scattering: A review", *Analytica Chimica Acta*, vol. 893, pp. 1-13, 2015. Available: 10.1016/j.aca.2015.04.010.

[40] A. Hakonen et al., "Detection of nerve gases using surface-enhanced Raman scattering substrates with high droplet adhesion", *Nanoscale*, vol. 8, no. 3, pp. 1305-1308, 2016.

[41] S. Veerabuthiran and A. Razdan, "LIDAR for Detection of Chemical and Biological Warfare Agents", *Defence Science Journal*, vol. 61, no. 3, pp. 241-250, 2011.

[42] A. K. Maini, Handbook of defence electronics and optronics: fundamentals, technologies and systems, John Wiley & Sons, 2018.

[43] V. A. Kovalev and W. E. Eichinger., *Elastic lidar: theory, practice, and analysis methods*,John Wiley & Sons, 2004.

[44] R. Robinson, T. Gardiner, F. Innocenti, P. Woods and M. Coleman, "Infrared differential absorption Lidar (DIAL) measurements of hydrocarbon emissions", *Journal of Environmental Monitoring*, vol. 13, no. 8, p. 2213, 2011.

[45] F. Innocenti, R. Robinson, T. Gardiner, A. Finlayson, and A. Connor. "Differential absorption lidar (DIAL) measurements of landfill methane emissions", *Remote Sensing*, vol. 9, no. 9, pp. 953, 2017.

[46] M. Milton, P. Woods, B. Jolliffe, N. Swann and T. McIlveen, "Measurements of toluene and other aromatic hydrocarbons by differential-absorption LIDAR in the near-ultraviolet", *Applied Physics B Photophysics and Laser Chemistry*, vol. 55, no. 1, pp. 41-45, 1992.

[47] M. Webber, M. Pushkarsky and C. Patel, "Optical detection of chemical warfare agents and toxic industrial chemicals: Simulation", *Journal of Applied Physics*, vol. 97, no. 11, p. 113101, 2005.

[48] L. Hoffland, R. Piffath and J. Bouck, "Spectral Signatures of Chemical Agents And Simulants", *Optical Engineering*, vol. 24, no. 6, 1985.

[49] S. W. Sharpe, T. J. Johnson, P. M. Chu, J. Kleimeyer, and B. Rowland, "Quantitative infrared spectra of vapor phase chemical agents", in *Chemical and Biological Sensing IV*, vol. 5085, pp. 19-27, International Society for Optics and Photonics, 2003.

[50] P. Geiko and A. Tikhomirov, "Remote measurement of chemical warfare agents by differential absorption CO2 lidar", *Optical Memory and Neural Networks*, vol. 20, no. 1, pp. 71-75, 2011.

[51] D. F. Flanigan, "Short history of remote sensing of chemical agents", in *Electro-Optical Technology for Remote Chemical Detection and Identification*, vol. 2763, pp. 2-17, International Society for Optics and Photonics, 1996.

[52] D. A. Leonard, T. A. Driscoll, and H. E. Sweeney. "Detection of organo phosphate vapors and liquids using a CO2 lidar", in *Optical Instruments for Weather Forecasting*, vol. 2832, pp. 20-31, International Society for Optics and Photonics, 1996.

[53] E. V. Degtiarev, A. R. Geiger, and R. D. Richmond. "Compact dual wavelength 3.30 to 3.47-um DIAL lidar", in *Chemical and Biological Sensing*, vol. 4036, pp. 229-235, International Society for Optics and Photonics, 2000.

[54] S. Chandra, M. E. Wager, B. L. Clayton, A. G. Geiser, T. H. Allik, J. L. Ahl, C. R. Miller et al., "2-um-pumped 8-12-um OPO source for remote chemical sensing", in *Chemical and Biological Sensing*, vol. 4036, pp. 200-208, International Society for Optics and Photonics, 2000.

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

[55] M. Kraus, F. Gebert, A. Walter, C. Pargmann and F. Duschek, "Online discrimination of chemical substances using standoff laser-induced fluorescence signals", *Journal of Chemometrics*, vol. 34, no. 2, 2020.
[56] Smiths Detection Technical Information [Online]. Available: https://www.smiths.com/-/media/files/sd_responderrci.pdf. [Accessed: 14-Apr-2020].
[57] H. Markert, J. Ring, N. Campbell, K. Grates, National Forensic Science Technology Ctr (NFSTC), and United States of America. "Evaluation of the Smiths Detection RespondeR™ RCI Raman Spectrometer-Technology Evaluation", 2010.
[58] "RespondeR RCI", *Cbrnetechindex.com*, 2020, [Online]. Available: https://www.cbrnetechindex.com/ Print/3524/ smiths-detection-inc/responder-rci. [Accessed: 19- Apr- 2020].
[59] "FirstDefender™ RM Chemical Identification System, Full Specifications ", *Thermofisher.com*, 2020. [Online]. Available: https://www.thermofisher.com/. [Accessed: 19- Apr- 2020].
[60] H. Markert, J. Ring, N. Campbell and K. Grates, National Forensic Science Technology Ctr (NFSTC), and United States of America. "Evaluation of the Thermo Scientific® FirstDefender RM™ Raman Spectrometer-Technology Evaluation", 2010.
[61] "FirstDefender brochure", *Thermofisher.com*, 2020. [Online]. Available: https://www.thermofisher.com/. [Accessed: 19- Apr- 2020].
[62] Raman Spectroscopy Instrument Maintenance Management, Support, Training and Repair - KD Analytical", *Kdanalytical.com*, 2020, [Online].Available: http://www.kdanalytical.com/instruments/technology/raman.aspx. [Accessed: 19- Apr- 2020].
[63] K. Gulati, V. Gambhir and M. Reddy, "Detection of Nitro-aromatic Compound in Soil and Sand using Time Gated Raman Spectroscopy", *Defence Science Journal*, vol. 67, no. 5, p. 588, 2017.
[64] K. Gulati, S. Gulia, T. Gambhir, N. Kumar, V. Gambhir and M. Reddy, "Standoff Detection and Identification of Explosives and Hazardous Chemicals in Simulated Real Field Scenario using Time Gated Raman Spectroscopy", *Defence Science Journal*, vol. 69, no. 4, pp. 342-347, 2019.
[65] S. Gupta, A. Ahmad, V. Gambhir, M. Reddy and D. Mehta, "Raman signal enhancement by multiple beam excitation and its application for the detection of chemicals", *Applied Physics Letters*, vol. 107, no. 9, p. 091114, 2015.
[66] "Our story — SEC Technologies", *SEC Technologies*, 2020, [Online]. Available: https://www.sec-technologies.com/our-technology, [Accessed: 19- Apr- 2020].
[67] M. van der Meer, J. A., and M. S. Nieuwenhuizen, "Observer report about field tests with a FALCON 4G stand-off detector by SEC Technologie", Rijswijk: TNO, 2016
[68] A. Puiu, L. Fiorani, O. Rosa, R. Borelli, M. Pistilli and A. Palucci, "Lidar/DIAL detection of acetone at 3.3 μm by a tunable OPO laser system", *Laser Physics*, vol. 24, no. 8, p. 085606, 2014.
[69] V. Mitev, S. Babichenko, R. Borelli, L. Fiorani, I. Grigorov, M. Nuvoli, A. Palucci et al., "Lidar extinction measurement in the mid infrared", in *20th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*, vol. 9292, p. 92923W, International Society for Optics and Photonics, 2014.
[70] V. Mitev, S. Babichenko, J. Bennes, R. Borelli, A. Dolfi-Bouteyre, L. Fiorani, L. Hespel et al, "Mid-IR DIAL for high-resolution mapping of explosive precursors", in *Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing IX*, vol. 8894, p. 88940S, International Society for Optics and Photonics, 2013.
[71] L. Fiorani, S. Babichenko, J. Bennes, R. Borelli, R. Chirico, A. Dolfi-Bouteyre, L. Hespel et al., "Lidar detection of explosive precursors", in *26th International Laser Radar Conference*, pp. 25-29. 2012.
[72] A. Boreysho et al., "Mobile multiwave lidar complexes", *Quantum Electronics*, vol. 35, no. 12, pp. 1167-1178, 2005.
[73] A. S. Boreysho, V. A. Volodenko, N. A. Gryaznov, E. R. Malamed, Y. N. Mendov, V. L. Moshkov, S. M. Pantaleev et al, "Mobile lidar complex for ecological monitoring of the atmosphere", in *Laser Optics 2003: High-Power Gas Lasers*, vol. 5479, pp. 177-186, International Society for Optics and Photonics, 2004.
[74] *Lsystems.ru*, 2020. [Online]. Available: http://www.lsystems.ru/en/products/14/206/. [Accessed: 19- April-2020].
[75] S. Yakovlev, S. Sadovnikov, O. Kharchenko and N. Kravtsova, "Remote Sensing of Atmospheric Methane with IR OPO Lidar System", *Atmosphere*, vol. 11, no. 1, p. 70, 2020.
[76] O. Romanovskii, S. Sadovnikov, O. Kharchenko and S. Yakovlev, "Remote analysis of methane content in the atmosphere by an IR DIAL lidar system in the 3300-3430-nm spectral range", *Atmospheric and Oceanic Optics*, no. 11, 2019.
[77] O. Romanovskii, S. Sadovnikov, O. Kharchenko and S. Yakovlev, "Development of Near/Mid IR differential absorption OPO lidar system for sensing of atmospheric gases", *Optics & Laser Technology*, vol. 116, pp. 43-47, 2019.
[78] O. Romanovskii, S. Sadovnikov, O. Kharchenko and S. Yakovlev, "Broadband IR Lidar for Gas Analysis of the Atmosphere", *Journal of Applied Spectroscopy*, vol. 85, no. 3, pp. 457-461, 2018.
[79] S.L. Jain, B.C. Arya, D.R. Nakra and A.K. Saha, "Differential absorption measurements of atmospheric minor constituents with a laser", *Indian Journal of Radio & Space Physics*, vol 15, pp. 29-31, 1986.
[80] R. Rossi, J. Ciparisse, A. Malizia, M. Gelfusa and P. Gaudio, "Multiwavelength differential absorption lidar to improve measurement accuracy: test with ammonia over a traffic area", *Applied Physics B*, vol. 124, no. 7, 2018.
[81] R. Rossi, J. Ciparisse, M. Gelfusa, A. Malizia and P. Gaudio, "The project TELEMACO: detection, identification and concentration measurements of hazardous chemical agents", *Journal of Instrumentation*, vol. 14, no. 03, pp. C03004-C03004, 2019.
[82] R. Rossi, D. Di Giovanni, A. Malizia and P. Gaudio, "Measurements of Vehicle Pollutants in a High-Traffic Urban Area by a Multiwavelength Dial Approach: Correlation Between Two Different Motor Vehicle Pollutants", *Atmosphere*, vol. 11, no. 4, p. 383, 2020.
[83] A. Razdan, S. Veerabuthiran, M. Jindal and R. Sharma, "Infrared differential absorption lidar for stand-off detection of chemical agents", *Pramana*, vol. 82, no. 2, pp. 385-389, 2014.
[84] S. Veerabuthiran, A. Razdan, M. Jindal, R. Sharma and V. Sagar, "Development of 3.0–3.45μm OPO laser based range resolved and hard-target differential absorption lidar for sensing of atmospheric methane", *Optics & Laser Technology*, vol. 73, pp. 1-5, 2015.

Mukesh Kumar Jindal obtained his B. Tech Degree from Kurukshetra University, India in Electronics and Instrumentation in 2001 and obtained his M. Tech Degree in Process Control from University of Delhi, India in the year 2006. He is currently working as Scientist-E in Laser Science & Technology Centre, Delhi, India. His current field of interest is to develop control electronics and software for LIDAR systems like Mie LIDAR, Differential Absorption LIDAR and Raman LIDAR etc.

Prof. Mainuddin received Bachelor's degree in Electronics and Communication Engineering from Jamia Millia Islamia in 1994. He received M.E. degree from the Department of Electronics and Communication Engineering, Delhi College of Engineering in 2003. He completed his Ph.D. in 2008 from Jamia Millia Islamia, New Delhi. He is presently working as a Professor at the department of Electronics and Communication Engineering, Jamia Millia Islamia, New Delhi. He worked as a Senior Research Scientist at Laser Science and Technology Centre, DRDO, New Delhi for more than 17 years. His research interests include antenna design, laser spectroscopy, Optoelectronics, Optical diagnostics, Sensors and Instrumentation and high power lasers, data communication, optical communication and computer

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3030672, IEEE Sensors Journal

networks etc. He has published more than 100 research articles in international journals/conferences.

Dr S. Veerabuthiran, Scientist F, completed his Ph D in the area of 'LIDAR for measurement of atmospheric aerosols and clouds' from Space Physics Laboratory/ISRO in 2003. He worked as a postdoctoral fellow in University of Sherbrooke, Sherbrooke, Canada. Currently, he is working in LASTEC, DRDO, Delhi on development of variants of LIDAR systems for various spectroscopic applications. He has more than 100 research articles in national & international journals and proceedings. Also, two monographs and one book chapter to his credit. He has filed two patents. His areas of research interest include development of LIDAR systems and laser spectroscopy based sensors for detection of biochemicals.

Anil Kumar Razdan, Scientist G, is Associate Director at Laser Science & Technology Centre (LASTEC), Delhi, India. After doing MSc in Physics, Dr Razdan obtained his PhD from Indian Institute of Technology Delhi in the area of Laser Spectroscopy. At LASTEC he has been working on development of high power lasers and laser based sensors for various applications. A life member of Indian Laser Association & Laser Spectroscopy Society of India, Dr Razdan has over 60 research publications to his credit both in national and international journals and has delivered several invited lectures both in India and abroad. He is recipient of DRDO Agni award for excellence in self reliance, as team member and Laboratory scientist of the year and technology group awards. He has co-authored two books titled *Laser and Bose Einstein Condensation Physics* (2010) from Narosa Publishers & *Advances in Laser Physics & Technology (2015)* from Cambridge University Press, India and has filed three patents. His areas of research interest include development of laser spectroscopy based sensors for remote sensing applications, development of high power laser systems, diagnostic techniques & adaptive optics. During last 10 years at LASTEC he has led 5 challenging projects related to laser based sensor development as Project Director. He is presently Head of LIDAR sensor technologies Division at LASTEC

1530-437X (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 06,2020 at 04:57:15 UTC from IEEE Xplore. Restrictions apply.

APPLIED FINANCE LETTERS

SPECIAL ISSUE 9.1, 2020

LIQUIDITY IN ASIAN FINANCIAL MARKETS: CROWDING OUT OR SPILLOVER EFFECT

SONAL THUKRAL[1], **RAHUL SIKKA**[1*]

1. Delhi Technological University, India

* Corresponding Author: Research Scholar, Delhi School of Management, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi – 110042, India. ✉ rahulsikka123@yahoo.co.in

Abstract

The paper attempts to explore the relationship between the stock market and the corporate bond market, with a focus on the inter-dependency of liquidity between the two markets. The study employs a panel dataset to assess the impact of stock market liquidity on the corporate bond market liquidity for the top five Asian economies (ranked by GDP) for the period 2008-2017. In contrast to a limited number of earlier studies that reported a spillover effect of liquidity among the markets for stock and government bonds, the results of the present study convey that an increase in stock market liquidity tends to eat up the liquidity of the corporate bonds, even after controlling for government bond yield and inflation rate changes. The findings indicate a crowding-out effect instead of a spillover effect, as indicated by related studies. The 'flight-to-quality' argument provides one possible explanation of liquidity moving away from one market to the other. This implies that if regulators' policies are focused on developing only one type of market, it may crowd out the liquidity and the development of the other market. The study suggests that the government focus more on the corporate bond market, which is yet to flourish in the Asian markets as compared to its stock market counterparts. The paper is one of the few attempts that focus on the corporate bond market and its liquidity and aims to ignite a debate on the possible linkages between the liquidity of the corporate bond market and the stock market.

Keywords: Liquidity, corporate bond market, stock market, turnover ratio, crowding out, spillover effect

1. Introduction

Financial Market liquidity has garnered increased interest in the past with close lens on the stock market. While focusing on the stock market, some studies have explored the effect of market liquidity on the prices and returns in the market (Amihud, 2002; Amihud and Mendelson, 1986; Jacoby et al., 2000; Jones, 2002; Jun et al., 2003; Pástor and Stambaugh, 2003). At the same time, others have explored factors that influence liquidity in the stock market (Correia and Amaral, 2014; Grossman and Miller, 1988; Hameed et al., 2010; Heflin et al., 2002; Kim et al., 2006). With regards to the bond market, the extant literature has analyzed the determinants of bond market development, but most of the work has focused on the size or capitalization of the market (Bhattacharyay, 2013; Eichengreen and Luengnaruemitchai, 2011; Maurya and Mishra, 2016; Mu et al., 2013; Smaoui et al., 2017; Teplova and Sokolova, 2018), and only a few studies examine the liquidity aspect of the bond market. Goyenko et al. (2011) and Gündüz et al. (2018) explore the factors that influence liquidity in the bond markets, while Lin et al. (2011) study the impact of liquidity on corporate bond returns. Further, the literature exploring the interlinkages of liquidity between the bond market and the stock market is almost missing as per our knowledge. Chordia et al. (2005) found liquidity shocks to be positively correlated across the two markets. In contrast, Goyenko and Ukhov (2009) found bond market

liquidity acting as a channel between monetary policy effects and the stock market liquidity, spilling over the money supply shocks from one market to another.

The limited studies that examine liquidity interlinkages between the bond and the stock markets overlook the corporate bond market and limit their focus to the treasury bond markets. Moving away from previous studies and in an attempt to fill this large gap, the present study attempts to explore the relationship between stock market liquidity and corporate bond market liquidity. Previous studies have explained stock market liquidity as a determinant of prices and returns of the stock market; however, our study differs by studying the stock market liquidity as a determinant of corporate bond market liquidity.

Further, the limited studies that focus on the liquidity of the corporate bond markets have concentrated on well-developed markets such as the United States (Bao et al., 2011; Chakravarty and Sarkar, 1999; Lin et al., 2011). Liquidity in the Asian corporate bond markets, which has been suffering (Sharma, 2001) has often escaped the attention of the researchers. To address this issue, our paper focuses on a cross country sample of Top five Asian economies.

A well-functioning corporate bond market makes it easier for corporations to avail credit and reduce the pressure on the banks leading to a reduction in the risk to the economy (Hakansson, 1999). A developed bond market promotes effective fundraising, which assists in stimulating economic activities (Hue and Trâm, 2019). Some studies establish a causal link between bond market development and economic growth (Fink et al., 2003; Pradhan et al., 2015), further emphasizing the importance of a bond market. Despite the importance of the bond market established in the literature, studies exploring its different aspects are relatively scarce, thereby heightening the importance to examine the functioning of these markets deeply. One of the significant reasons for scant literature can be attributed to the lack of trading level data as pointed out by Harrison (2002).

Asian economies have been growing at an impressive rate. Currently, the share of Asian economies in the global output is at 38%, rising from 26% in the 2000s and 2020, Asian economies are expected to surpass rest of the world in terms of Purchasing Power Parity (Romei and Reed, 2019). Asian corporate bond markets have been growing in size, measured by the number of outstanding issues, for the past few years, and in March 2018, reached a valuation of $5.3 trillion (CARE Ratings Ltd., 2018). However, when we compare it with the US market, which stands at more than $8.8 trillion, Asian countries seem to have lagged. The vast potential in the Asian markets provides impetus to explore the corporate bond market scenario in Asia.

To study the effect that stock market liquidity has on the liquidity in the corporate bond market, the present study employs a panel of top five Asian economies ranked by their GDP for the period 2008 to 2017. Utilizing insights from the extant literature, our study also controls for two other variables that could have a possible effect on the corporate bond market liquidity – government bond yield and inflation rate in the economy. The results of the study show a negative relationship between the corporate bond market liquidity and the stock market liquidity. The negative relationship observed in our results points towards a crowding-out effect between the two markets explaining that higher liquidity in the stock market seems to eat up the liquidity in the corporate bond market. The relationship remains significant, even after controlling for government bond yields and inflation rates in the economy.

The contribution of the present study lies in its attempt to draw the attention of the researchers to the highly overlooked area, the corporate bond market liquidity. Existing literature assesses the liquidity linkages between the treasury bond markets and stock markets; however, our study explores the linkages of liquidity pertaining to the corporate bond market and the stock market. The crowding-out effect observed in the study provides an interesting aspect of the dynamic relationship between the two markets. Previous studies, ostensibly similar to our present study that examines the linkages between the stock market and bond market liquidity (Chordia et al., 2005; Goyenko and Ukhov, 2009) have found evidence of a spillover effect rather than the crowding-out

effect. The present study also observes that stock market liquidity in our overall sample surpasses the liquidity in the corporate bond market, possibly owing to the lucrativeness of the stock market for investors. The lucrativeness can be attributed to the efficiencies brought in the stock markets by concentrated efforts of the government to uplift these markets. Thus, the crowding-out effect of stock market development emphasizes the need of developing policies which not only targets one market (stock market, in this case) but similar and simultaneous efforts are required to uplift other markets (corporate bond markets, in the present case) to prevent the flight of trades to "quality" markets.

2. Bond Market Trends in Asia

Asian bond markets are growing in size; however, they still lag behind the western countries, primarily the USA. In addition to this, the contribution of each Asian economy towards the total size of the Asian market is quite skewed, with only China contributing close to 51% of the total market size. China's bond market is ranked among the top bond markets of the world in terms of size; however, the market is dominated by the Government or treasury bonds and the interbank market, with only about 21% being contributed by the corporate bonds (ACRA, 2019). The Republic of Korea follows China in terms of bond market size, which is followed by Japan and India, respectively. However, Indonesia, when compared to its Asian peers' lags in size (CARE Ratings Ltd., 2018). According to the same report, however, we see the turnover ratio of the Indonesian corporate bond market is the highest in 2018, whereas Japan and China lag at the bottom. The size (measured by the stock market capitalization) of the stock markets in the above countries far exceed the size of their corporate bond markets. Japan's exchange alone has a capitalization of more than $5 trillion (Statista, 2019), compared to their bond market size of only $700 billion. Even in India's case, we observe a similar difference, while their bond market size was around $400 billion, the market capitalization of BSE in 2018 was close $2 trillion. These differences shed light on the lack of development of the corporate bond markets in comparison to the stock markets in the Asian economies.

3. Background of the Study

3.1 Stock market liquidity and Bond market Liquidity

Limited studies in the past explore the interlinkages of the stock market and bond market, especially the liquidity aspect. Volatility in the market returns primarily is positively correlated among the two markets (Fleming et al., 1998). On the other hand, there have also been studies in the past highlighting the impact of 'flight-to-quality', where the funds have moved from the stock market to the bond market on the assumption that bond markets are a safer option, especially in times of uncertainty (Bayraci et al., 2018; Connolly et al., 2005). There is consensus among the researchers about the negative return correlations among the two markets; however, when comparing the liquidity of the bond and stock markets, results have been different. While studying the relationship between the liquidity of the stock market and that of the treasury bond market, previous works have found a positive association between the two (Chordia et al., 2005). While studying the impact of monetary policy on illiquidity, Goyenko and Ukhov (2009) observed a spillover effect where a liquidity shock in the bond market was subsequently transmitted to the stock market. These studies offer insights into the liquidity relationship across markets; however, the lack of studies examining the corporate bond liquidity leaves a gap to be explored.

Liquidity becomes a critical component of any financial market, and with meagre literature that examines the liquidity of the corporate bond market, the present study attempts to bridge this gap. Studying the liquidity element of the financial markets has important implications. The trading activity has been found to have a positive correlation with the returns of the markets (Edelen and

Warner, 2001). Level of trading activity often signals the level of confidence that investors have regarding the market conditions. Markets with higher liquidity experience lower transaction costs and enhanced market efficiency (Chordia et al., 2008). Since liquidity act as a market signal, lower liquidity in the stock market would reduce investor confidence and investors could move onto the relatively stable corporate bond market, resulting in higher trading activity in the corporate bonds.

H1: Stock market liquidity is negatively related to corporate bond market liquidity.

3.2 Government Bonds Yields and Liquidity of the Corporate Bond Market

While studying the liquidity effect on yield spreads, Chen et al. (2007) found an inverse relationship. They found higher yield spreads to be associated with more illiquid bonds, indicating that liquidity is priced in the market. These results were further backed by Favero et al. (2010), who found liquidity risk results in yield differentials in the eurozone sovereign bonds. A higher yield in the government bond market would suggest a fall in liquidity. Amihud and Mendelson (1991) highlight the implications of high yield and low liquidity in the market. Issuers would not prefer a less liquid market as they would not be able to sell securities at a desirable price. Also, they propose that investors demand less of illiquid instruments. The insights from these works suggest that a higher yield in the treasury bond market would negatively affect its desirability and would provide investors and issuers, both, an incentive to explore other markets. The negative association between the treasury bonds and corporate bonds have been explored in the work of Duffee (1998). The author found a significant negative relationship between the yields of the two markets. Results from the study show a fall in the yields of the investment-grade corporate bonds as the yields of the treasury market rises. The discussion suggests that a higher yield in the government bonds will result in a drop in corporate bond yields. Consequently, this would be associated with higher liquidity in the corporate bond market as it becomes more desirable for both investors and issuers, based on the implications provided by Amihud and Mendelson (1991). Therefore, we hypothesize a positive relationship between government bond yields and liquidity in the corporate bond market.

H2: Government bond yield is positively related to corporate bond market liquidity.

3.3. Inflation rate and Corporate Bond market liquidity

Supply of money in the market has a positive impact on liquidity (Ariff et al., 2012). Extra supply of money in the market pushes the interest rate down, through the 'liquidity effect' (Christiano, 1991); fall in interest rates boosts the economic activity and investment, resulting in a positive effect on the trading activity in the financial markets. Inflation in an economy is a significant result of money supply shocks, with a strong positive association between the level of money and consumer prices (Browne and Cronin, 2010) and a rise in inflation signals quantitative easing. A quantitative easing strategy tends to reduce the price frictions and provides a boost to trading activity, thus having a positive impact on market liquidity (Christensen and Gillan, 2013). Thus, higher inflation rates should result in higher liquidity in the corporate bond market.

H3: Inflation rates are positively related to corporate bond market liquidity.

4. Data and Methodology

The study first presents the trends in turnover ratios in the Asian economies, followed by descriptive statistics. Next, to study the impact of stock market liquidity on the corporate bond market liquidity, we employed a panel dataset. Our sample consists of the top five Asian economies, ranked by GDP (IMF, 2019b) viz., China, India, Japan, Republic of Korea and Indonesia for the period 2008-2017. The model includes bond turnover ratio as the dependent variable, stock turnover ratio as independent variable and government bond yield and inflation rate as control variables. The data for all the variables have been collected on an annual basis. The proxies for variables and the respective

sources of data collection have been explained in the following sections and summarized in Table 1.

4.1. Measures of Liquidity

To study the relationship between the liquidity of the stock market and the corporate bond market, we utilize the turnover ratios of the two markets for our sample Asian countries. The turnover ratio is calculated using the definition given by Asian Bonds Online, as the ratio of the total value of securities traded in the market in a year and the average market capitalization or the number of outstanding issues in the market in that year. The bond market development indicators given by World Bank lists the turnover ratio as an essential variable for testing the efficiency of the market (The World Bank, 2006). The turnover ratios have also been previously utilized by researchers to study market liquidity (Chan et al., 2007; Jun et al., 2003; Sarr and Lybek, 2002).

4.1.1 Corporate bond market liquidity (btr) –The liquidity of the corporate bond market in an economy is proxied by the corporate bond turnover ratio (btr). Corporate bond turnover ratios of the countries in our sample except for India has been extracted from the data portal of Asian Bonds Online (AsianBondsOnline, 2019). For India, the amount of trading in the corporate bond market and the number of outstanding issues in the market have been extracted from the Securities Exchange Board of India's (SEBI) 'Reports and Statistics' on corporate bonds (SEBI, 2019). For maintaining consistency, the bond turnover ratios for the same period for India have been calculated using the same definition as Asian Bonds Online.

4.1.2 Stock Market Liquidity (stock turnover) – Stock market liquidity has been proxied by the stock turnover ratio. The stock turnover ratio for all the economies in the dataset has been extracted from The World Bank's 'Global Financial Development Database' (The World Bank, 2019).

4.2. Control variables

4.2.1 Government Bond yields (gbond) – Yields of the 10-year Treasury bond have been taken as a proxy. The data for government bond yields have been extracted from the International Monetary Fund's (IMF) International financial statistics database for all the economies in our dataset (IMF, 2019a).

4.2.2 Inflation Rate (infperc) – The second control variable is the inflation rate, We extract inflation rates from the IMF's World Economic Outlook Database, October 2019 edition (IMF, 2019b).

4.3. Model Specification

Table 1: Description of Variables

Variable	Symbol	Proxy	Role	Data Source
Corporate Bond liquidity	btr	Corporate bond turnover ratio	Dependent Variable	Extracted from Asian bonds online and SEBI – Reports and Statistics
Stock Market Liquidity	Stock-turnover	Stock turnover ratio	Independent Variable	World Bank's Global Financial Development Database
Government/Treasury Bond market	gbond	Annual yield of the 10-year treasury bond.	Control Variables	IMF's International Financial Statistics
Inflation rate	infperc	Annual inflation rate		IMF's World Economic Output database

Given the panel structure of our dataset and the need to test the impact of stock market liquidity on bond market liquidity, we employ panel regression. Controlling for heteroscedasticity and autocorrelation by using Arellano specification (Arellano, 1987), the model was checked for the

presence of fixed effects or random effects using the Hausman Test (Hausman and Taylor, 1981). The results suggested for fixed effects, and hence the same was applied. The following specification model was used:

$$\log(btr)_{ti} = \alpha + \beta_1 * \log(stockturnover)_{ti} + \beta_2 * \log(gbond)_{ti} + \beta_3 * \log(infperc)_{ti} + \mu_{ti.} \quad (1)$$

5. Results and Analysis

Summary statistics for the variables are reported in Table 2. The mean turnover ratio of the corporate bond market is found to be lower when compared with the stock market, and similar observation is noted for the median values. Lower liquidity in the corporate bond market across our sample indicates more activeness of the investors in the stock. The difference in the mean liquidity could point towards a lack in the development of the corporate bond market and shows how the stock market is preferred over the former. When comparing the individual countries in the sample (refer Table 3) based on the liquidity of the markets, we observe China's mean turnover ratio for the bond market is close to 2.848, much higher than all other countries in the sample, with the second-highest turnover being that of the Indonesian bond market at only 0.8. India, in comparison, lags with a mean turnover ratio of only 0.6 in the sample period. The corporate bond market of China has seen a rapid fall in its turnover ratio in the past ten years. The average turnover ratio for the first five years is found to be around 4 for the Chinese bond market, whereas for the next five years, it falls to approximately 1. Surprisingly, in the same period, the size of the Chinese corporate bond market has seen a significant expansion, with its size in 2017 being eight times the size in 2008 (AsianBondsOnline, 2019). This observation indicates that size may not be a strong determinant of liquidity, and merely increasing the issuance will not guarantee a rise in liquidity. Figure 1 and 2 display the trends of bond turnover ratio and stock turnover ratio, respectively, of individual countries in our sample. Indian corporate bond market has not seen much variability in its turnover ratio, and it has remained almost stagnant in-sample period.

In contrast, China experienced a considerable dip in the bond turnover ratio. Rest of the economies also do not observe much variability in the bond turnover ratios. Studies that focus on the development of the Indian corporate bond market associate lack of market makers as a reason for low liquidity (Banerji et al., 2012). Indian Insurance companies and pension funds, both of which are huge players in the market and act as market makers, are restricted by mandates specified by the regulators and are permitted to invest only in bonds recognized as 'investment' grade (PFRDA, 2014; IRDAI, 2016).

Table 2: Descriptive Statistics- Overall Sample

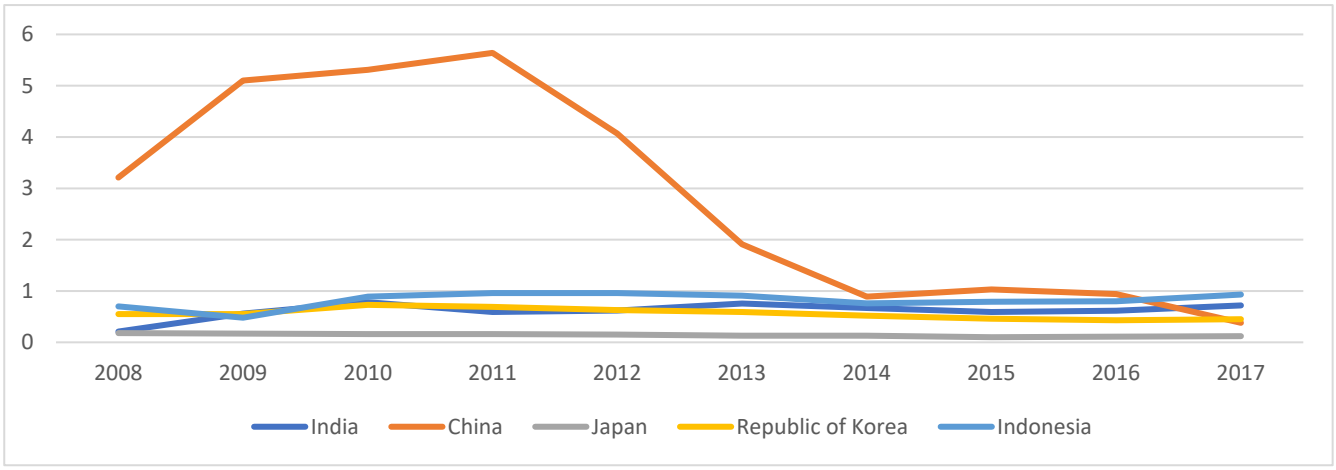
	Bond Turnover Ratio	Stock turnover Ratio	Government Bond Yields	Inflation rates
Mean	0.996	1.219	4.814	3.752
Median	0.624	1.133	4.178	3.12
Standard Deviation	1.31	0.938	3.082	3.302
Sample Variance	1.717	0.881	9.498	10.903
Minimum	0.1	0.196	-0.066	-1.35
Maximum	5.64	5.569	12.2	12.314
Count	50	50	50	50

Table 1: Descriptive Statistics - bond turnover ratio - Country-wise for the period 2008-2017

India		China		Japan	
Mean	0.61	Mean	2.85	Mean	0.14
Median	0.62	Median	2.56	Median	0.14
Standard Deviation	0.16	Standard Deviation	2.06	Standard Deviation	0.03
Minimum	0.21	Minimum	0.38	Minimum	0.10
Maximum	0.78	Maximum	5.64	Maximum	0.18

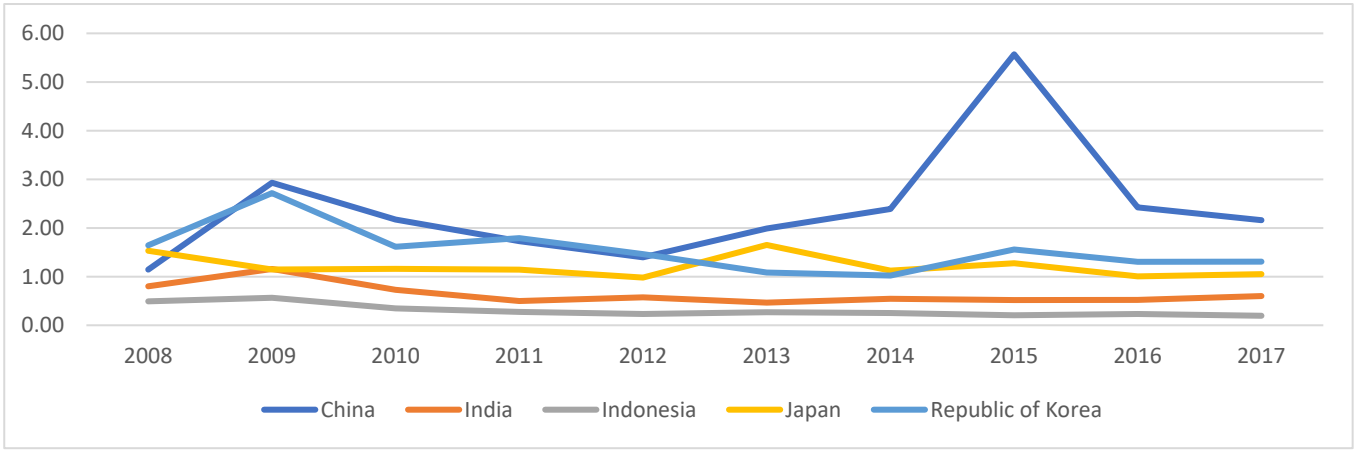
Republic of Korea		Indonesia	
Mean	0.56	Mean	0.82
Median	0.55	Median	0.85
Standard Deviation	0.10	Standard Deviation	0.15
Minimum	0.43	Minimum	0.48
Maximum	0.73	Maximum	0.96

Figure 1 - Bond Turnover Ratios of the top five Asian Economies



Source: Asian Bonds Online and SEBI – Reports and Statistics.

Figure 2 - Stock Turnover Ratios of the top five Asian Economies



Source: World Bank – Global Financial Development Database

Next, we applied the fixed effects model to test the impact of stock turnover ratio on bond turnover ratio; results are summarized in Table 4. First, we ran the model, including only the dependent variable and the independent variable (omitting the control variables). Results are reported in column 2 of Table 4. It was found that the stock turnover ratio had a negative coefficient and was significant, indicating that a rise in the liquidity of the stock market would result in a fall in liquidity in the corporate bond market. However, the model resulted in a low coefficient of determination and to improve the explanatory power of the model; we included the two control variables.

Table 2: Fixed effects model results

Variables (1)	**Model with only independent variable** (2)	**Model with control variables and robust standard errors** (3)
log(stockturnover)	0.0426* (-0.4573)	0.000*** (-0.5833)
log(gbond)		0.7112 (0.0467)
log(infperc)		0.0467 (0.3237)
R^2	0.0902	0.274
F statistic (p-value)	0.0426*	0.0086**

Note: Numbers in the parenthesis indicate respective variables estimates. ***, **, * indicate significance at 0.1%, 1% and 5% levels, respectively.

With the inclusion of control variables viz. government bond yield and inflation rates, also controlling for autocorrelation and heteroscedasticity by calculating robust standard errors (reported in column 3 of Table 4), results indicate that stock turnover ratio is still highly significant and is negatively related to the bond turnover ratio. The observed negative relationship between the liquidity of the corporate bond market and the stock market is an interesting finding of the study. Whereas previous studies found stock market liquidity and treasury bond market liquidity to be moving in tandem, our results show liquidity to moving away from one market to another. The argument of 'flight-to-quality' has been previously used to explain the negative relationship between price movements in the bond market and the stock market. According to the argument, in times of market turbulence, investors prefer a stable bond market over the stock market (Ilmanen, 2003; Bayraci, Demiralay and Gencer, 2018). Fang et al. (2017) found similar results, where the level of certainty in the economy affected the correlation in the price movements among the bond and stock markets. We can extend this explanation of negative price relationship between the two markets to the negative relationship of liquidity, we observed in our results. In a stable market scenario, the trading in equity rises whereas a negative sentiment would transfer the trading activity towards the bond market, which is relatively less risky and investors look to diversify their risk (Connolly et al., 2005). Stock market liquidity is regarded as one of the predictors of the business cycle and the future state of the economy (Næs et al., 2011). Higher liquidity in the stock market is associated with a favourable economic situation and vice-versa. These observations further provide support to our results, as lower liquidity in stock market signals a relatively volatile economic situation and thus, the transfer of liquidity to the more stable bond markets. Also, another interesting aspect of the discussion is the relatively lower liquidity of the corporate bond markets. The regulations in these markets restrict particular types of investments in non-investment grade bonds which reduces the ability of institutions to trade in the market effectively and further diminishes their market-making capabilities. Due to such regulations, much of the liquidity that could have been otherwise in the bond market is being crowded out by the liquidity and hence the attractiveness of the stock market.

Our results also show a positive relationship between inflation and bond turnover ratio; however, the relationship is not found to be significant. Inflation has often been linked with the increase in money supply (Orphanides and Solow, 1990). As the money supply improves, the funds allocated to various assets improve, thereby increasing the trading activity. Money supply and liquidity have been reported in the past as having a strong positive relationship (Ariff et al., 2012), and a similar effect is

reported in our results. Government bond yield is found to be positive but not significantly related to bond turnover ratio. Duffee (1998) found a negative relationship between the yields of the treasury bond market and the corporate bond market, as the yields in the treasury bond market increases, yields in the corporate bond market fall. Lower yields in the corporate bond market signal higher liquidity (Chen et al., 2007). In short, results from our study suggest that corporate bond market liquidity is not only affected by the liquidity in the stock markets, but their relationship is found to be inverse in nature. Higher liquidity in the stock market tends to crowd out the liquidity in the corporate bond market.

6. Conclusion

In our paper, we attempted to explore the nature of the relationship between the stock market liquidity and the corporate bond market liquidity. The study included a sample of the top five Asian economies for the period 2008-2017. Studies on the liquidity linkages between the bond market and stock market have been limited in the literature, and most of the research focuses on the Treasury bond markets. Our study expands on the scant literature on the corporate bond market liquidity and its relationship with the stock market. Our analysis begins with studying the levels and trends of liquidity of the stock and the corporate bond market in the countries included in our sample. It was observed that the overall corporate bond market liquidity had lagged that of the stock market. Another interesting observation made in the paper is the difference between the size of the corporate bond market and the respective liquidity. China, despite having a considerable market size, saw a fall in liquidity in the last few years.

On the other hand, Indonesia, which has the smallest bond market in the sample, had overall liquidity as the second highest in the sample, ahead of the larger markets of India and Japan. These observations signal that size alone may not be a determinant of liquidity. Further, Fixed effects model was employed to assess the impact of stock market liquidity on the corporate bond market liquidity. Even after controlling for government bond yields and inflation rate changes, a significant inverse relationship was found among the liquidity of the two markets, indicating that the liquidity of the two markets moves in the opposite direction. Thus, the results point towards a crowding-out effect, which is not explicitly explored in extant literature. Close to our present paper, the relationship between treasury bond market and the stock market has been explored in the past (Chordia et al., 2005; Goyenko and Ukhov, 2009), which has indicated a spillover effect of liquidity in the treasury bond market on liquidity in the stock market. Under this spillover effect, a liquidity shock in one market is followed by a liquidity shock in the other; however, our results indicate that a rise in liquidity in one market eats up the liquidity in the other market. The negative relationship between the liquidity of the two markets can be explained through the 'flight-to-quality' argument (Connolly et al., 2005; Bayraci et al., 2018). According to the 'flight-to-quality' explanation, any disturbance in the stock market may result in the movement of funds from the stock market to a relatively less risky market, the bond market, thereby increasing the trading activity (liquidity) in the latter market.

The fact can also support the crowding out effect of the stock market liquidity on the corporate bond market liquidity, that the Asian economies in our sample have a more liquid stock market compared to the corporate bond market, as observed through the descriptive statistics. The plausible reason for the same seems to be a lack of regulatory policies in making the bond market as lucrative as the stock market. For instance, the market regulators in India, restrict financial corporations, like the insurance companies, to invest in non-investment grade bonds, thus, hindering their ability to act as market makers (Banerji et al., 2012).

The findings of the present study support the opposing relationship between the liquidities of the stock and the corporate bond market and provide evidence of a crowding-out effect. This calls for a more concentrated effort by the regulators to develop not only one market but to have simultaneous developments in other markets as well. The paper aims to provide insights into the relatively unexplored corporate bond market liquidity, and our results bring attention to the need for

simultaneous efforts to promote liquidity in both the markets. Policies that focus on liquidity and not just the size of the market are essential, calling for the reforms needed to boost market activity, particularly in the case of the corporate bond market. The paper aims to ignite the debate on the interlinkages between the stock market and the corporate bond market.

7. Limitations and Future Research Scope

The paper leaves some scope for future research. A high-frequency data of turnover ratio of the stock and corporate bond markets, besides extending the analysis to other economies for an extended sample period would shed light on more delicate nuances of the inter-linkages between the two markets. The present paper skips the impact of various economic conditions on the inter-linkages between the liquidity of the two markets and is left for future researchers to explore. A comparison of developed and developing economies may provide further insights. One of the reasons for low liquidity in the corporate bond market can also be attributed to the attractiveness of the treasury bond market. Researchers may explore if a crowding-out effect exists between the corporate bond market and the treasury bond market. A model that incorporates the impact of macroeconomic factors in explaining the liquidity impact of the stock market on the corporate bond market may also be explored with a specific focus on the regulatory environment in the markets.

References

ACRA (2019) A brief introduction to China's bond market. doi: https://www.acra-ratings.com/research/1116.

Amihud, Y. (2002) 'Illiquidity and stock returns: cross-section and time-series effects', Journal of Financial Markets, 5(1), pp. 31–56. doi: https://doi.org/10.1016/S1386-4181(01)00024-6.

Amihud, Y. and Mendelson, H. (1986) 'Asset pricing and the bid-ask spread', Journal of Financial Economics, 17(2), pp. 223–249. doi: https://doi.org/10.1016/0304-405X(86)90065-6.

Amihud, Y. and Mendelson, H. (1991) 'Liquidity, Asset Prices and Financial Policy', Financial Analysts Journal. Routledge, 47(6), pp. 56–66. doi: 10.2469/faj.v47.n6.56.

Arellano, M. (1987) 'Practitioners'corner: Computing robust standard errors for within-groups estimators', Oxford Bulletin of Economics and Statistics. Wiley Online Library, 49(4), pp. 431–434.

Ariff, M., Chung, T. and Shamsher (2012) 'Money supply, interest rate, liquidity and share prices: A test of their linkage', Global Finance Journal, 23(3), pp. 202–220. doi: https://doi.org/10.1016/j.gfj.2012.10.005.

AsianBondsOnline (2019) Asian Development Bank, Asian Development Bank. Available at: https://asianbondsonline.adb.org/data-portal/.

Banerji, S. et al. (2012) New Thinking on Corporate Bond Market in India New thinking on corporate bond market in India. 106. New Delhi.

Bao, J., Pan, J. U. N. and Wang, J. (2011) 'The Illiquidity of Corporate Bonds', The Journal of Finance. Wiley Online Library, 66(3), pp. 911–946. doi: 10.1111/j.1540-6261.2011.01655.x.

Bayraci, S., Demiralay, S. and Gencer, H. G. (2018) 'Stock and Bond Co-Movements and Flight-to-Quality in G7 Countries: A Time Frequency Analysis', Bulletin of Economic Research. John Wiley & Sons, Ltd (10.1111), 70(1), pp. E29–E49. doi: 10.1111/boer.12118.

Bhattacharyay, B. N. (2013) 'Determinants of bond market development in Asia', Journal of Asian Economics, 24, pp. 124–137. doi: 10.2139/ssrn.1898047.

Browne, F. and Cronin, D. (2010) 'Commodity prices, money and inflation', Journal of Economics and Business, 62(4), pp. 331–345. doi: https://doi.org/10.1016/j.jeconbus.2010.02.003.

CARE Ratings Ltd. (2018) Asian Corporate Bond Market – A comparison with India. Available at: http://www.careratings.com/upload/NewsFiles/Economics/Corporate Bonds - A Comparison with India.pdf.

Chakravarty, S. and Sarkar, A. (1999) 'Liquidity in US fixed income markets: A comparison of the bid-ask spread in corporate, government and municipal bond markets', FRB of New York Staff Report, (73).

Chan, E., Ahmad, M. F. H. and Wooldridge, P. (2007) 'Liquidity in an emerging bond market: a case study of corporate bonds in Malaysia', Bank of International Settlements (BIS) Working Paper Series.

Chen, L., Lesmond, D. A. and Wei, J. (2007) 'Corporate Yield Spreads and Bond Liquidity', The Journal of Finance. John Wiley & Sons, Ltd (10.1111), 62(1), pp. 119–149. doi: 10.1111/j.1540-6261.2007.01203.x.

Chordia, T., Roll, R. and Subrahmanyam, A. (2008) 'Liquidity and market efficiency', Journal of Financial Economics, 87(2), pp. 249–268. doi: 10.1016/j.jfineco.2007.03.005.

Chordia, T., Sarkar, A. and Subrahmanyam, A. (2005) 'An Empirical Analysis of Stock and Bond Market Liquidity', The Review of Financial Studies. New York, NY: [Oxford University Press, Society for Financial Studies], 18(1), pp. 85–129. Available at: http://www.jstor.org/stable/3598068.

Christensen, J. H. E. and Gillan, J. M. (2013) Does Quantitative Easing Affect Market Liquidity? doi: https://doi.org/10.24148/wp2013-26.

Christiano, L. J. (1991) 'Modeling the liquidity effect of a money shock', Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis. Federal Reserve Bank of Minneapolis, 15(1), p. 3.

Connolly, R., Stivers, C. and Sun, L. (2005) 'Stock Market Uncertainty and the Stock-Bond Return Relation', The Journal of Financial and Quantitative Analysis. Cambridge University Press, 40(1), pp. 161–194. Available at: http://www.jstor.org/stable/27647190.

Correia, L. and Amaral, H. (2014) 'Determinants of Market Liquidity of Shares Traded on the BM&FBOVESPA', Brazilian Business Review, 11(6), pp. 75–97. doi: 10.15728/bbr.2014.11.6.4.

Duffee, G. R. (1998) 'The Relation Between Treasury Yields and Corporate Bond Yield Spreads', The Journal of Finance. John Wiley & Sons, Ltd (10.1111), 53(6), pp. 2225–2241. doi: 10.1111/0022-1082.00089.

Edelen, R. M. and Warner, J. B. (2001) 'Aggregate price effects of institutional trading: a study of mutual fund flow and market returns', Journal of Financial Economics, 59(2), pp. 195–220. doi: https://doi.org/10.1016/S0304-405X(00)00085-4.

Eichengreen, B. and Luengnaruemitchai, P. (2011) 'Why Doesn't Asia Have Bigger Bond Markets?', SSRN Electronic Journal. Cambridge, MA. doi: 10.2139/ssrn.1009056.

Fang, L., Yu, H. and Li, L. (2017) 'The effect of economic policy uncertainty on the long-term correlation between U.S. stock and bond markets', Economic Modelling, 66, pp. 139–145. doi: https://doi.org/10.1016/j.econmod.2017.06.007.

Favero, C., Pagano, M. and von Thadden, E. L. (2010) 'How Does Liquidity Affect Government Bond Yields?', Journal of Financial and Quantitative Analysis. 2009/11/26. Cambridge University Press, 45(1), pp. 107–134. doi: DOI: 10.1017/S0022109009990494.

Fink, G., Haiss, P. R. and Hristoforova, S. (2003) 'Bond Markets and Economic Growth', Europainstitut Working Paper No. 49. Available at SSRN Electronic Journal. Elsevier BV. doi: 10.2139/ssrn.1003763.

Fleming, J., Kirby, C. and Ostdiek, B. (1998) 'Information and volatility linkages in the stock, bond, and money markets.', Journal of Financial Economics, 49(1), pp. 111–137. doi: https://doi.org/10.1016/S0304-405X(98)00019-1.

Goyenko, R., Subrahmanyam, A. and Ukhov, A. (2011) 'The Term Structure of Bond Market Liquidity and Its Implications for Expected Bond Returns', Journal of Financial and Quantitative Analysis. 2010/11/10. Cambridge University Press, 46(1), pp. 111–139. doi: DOI: 10.1017/S0022109010000700.

Goyenko, R. and Ukhov, A. (2009) 'Stock and Bond Market Liquidity: A Long-Run Empirical Analysis', The Journal of Financial and Quantitative Analysis. [Cambridge University Press, University of Washington School of Business Administration], 44(1), pp. 189–212. Available at: http://www.jstor.org/stable/40505920.

Grossman, S. J. and Miller, M. H. (1988) 'Liquidity and Market Structure', The Journal of Finance, 43(3), p. 617. doi: 10.2307/2328186.

Gündüz, Y. et al. (2018) 'Lighting up the Dark: Liquidity in the German Corporate Bond Market', SSRN Electronic Journal. doi: 10.2139/ssrn.3253000.

Hakansson, N. H. (1999) 'The Role of a Corporate Bond Market in an Economy--and in Avoiding Crises', SSRN Electronic Journal. Elsevier BV. doi: 10.2139/ssrn.171405.

Hameed, A., Kang, W. and Viswanathan, S. (2010) 'Stock market declines and liquidity', Journal of Finance, 65(1), pp. 257–293. doi: 10.1111/j.1540-6261.2009.01529.x.

Harrison, P. (2002) 'The impact of market liquidity in times of stress on the corporate bond market: pricing, trading, and the availability of funds during heightened illiquidity', in Third Joint Central Bank Research Conference on Risk Measurement and Systemic Risk. Bank of International Settlements (BIS).

Hausman, J. A. and Taylor, W. E. (1981) 'Panel data and unobservable individual effects', Econometrica: Journal of the Econometric Society. JSTOR, pp. 1377–1398.

Heflin, F., Shaw, K. W. and Wild, J. J. (2002) 'Disclosure quality and market liquidity', Working Paper, Robert H. Smith School of Business, College Park, MD.

Hue, N. T. M. and Trâm, N. T. N. (2019) 'Bond Market Development and Economic Growth : An Empirical Study in Asia', Journal of Asian Review of Public Affairs and Policy, pp. 1–12.

Ilmanen, A. (2003) 'Stock-bond correlations', The Journal of Fixed Income. Institutional Investor Journals Umbrella, 13(2), pp. 55–66. doi: https://doi.org/10.3905/jfi.2003.319353.

IMF (2019a) International Financial Statistics, International Monetary Fund. Available at: http://data.imf.org/?sk=4C514D48-B6BA-49ED-8AB9-52B0C1A0179B.

IMF (2019b) World Economic Database.

IRDAI (2016) IRDAI (Investment) Regulations. India. Available at: https://www.irdai.gov.in/ADMINCMS/cms/frmGeneral_Layout.aspx?page=PageNo2934&flag=1.

Jacoby, G., Fowler, D. J. and Gottesman, A. A. (2000) 'The capital asset pricing model and the liquidity effect: A theoretical approach', Journal of Financial Markets, 3(1), pp. 69–81. doi: https://doi.org/10.1016/S1386-4181(99)00013-0.

Jones, C. (2002) 'A Century of Stock Market Liquidity and Trading Costs', SSRN Electronic Journal. doi: 10.2139/ssrn.313681.

Jun, S. G., Marathe, A. and Shawky, H. A. (2003) 'Liquidity and stock returns in emerging equity markets', Emerging Markets Review, 4(1), pp. 1–24. doi: https://doi.org/10.1016/S1566-0141(02)00060-2.

Kim, J. C., Jain, P. K. and Rezaee, Z. (2006) 'Trends and Determinants of Market Liquidity in the Pre- and Post-Sarbanes-Oxley Act Periods', in 14th Annual Conference on Financial Economics and Accounting (FEA). doi: dx.doi.org/10.2139/ssrn.488142.

Lin, H., Wang, J. and Wu, C. (2011) 'Liquidity risk and expected corporate bond returns', Journal of Financial Economics, 99(3), pp. 628–650. doi: https://doi.org/10.1016/j.jfineco.2010.10.004.

Maurya, A. K. and Mishra, O. N. (2016) 'The Macro-economic Determinants of Corporate Bond Market in India', PACIFIC BUSINESS REVIEW INTERNATIONAL, 8(12), pp. 163–176.

Mu, Y., Phelps, P. and Stotsky, J. G. (2013) 'Bond markets in Africa', Review of Development Finance, 3(3), pp. 121–135. doi: 10.1016/j.rdf.2013.07.001.

Næs, R., Skjeltorp, J. A. and Ødegaard, B. A. (2011) 'Stock Market Liquidity and the Business Cycle', Journal of Finance. John Wiley & Sons, Ltd (10.1111), 66(1), pp. 139–176. doi: 10.1111/j.1540-6261.2010.01628.x.

Orphanides, A. and Solow, R. M. B. T. H. of M. E. (1990) 'Chapter 6 Money, inflation and growth', in. Elsevier, pp. 223–261. doi: https://doi.org/10.1016/S1573-4498(05)80009-8.

Pástor, Ľ. and Stambaugh, R. F. (2003) 'Liquidity Risk and Expected Stock Returns', Journal of Political Economy. The University of Chicago Press, 111(3), pp. 642–685. doi: 10.1086/374184.

PFRDA (2014) 'Revision of Investment Guidelines for NPS Schemes'. New Delhi: Penson Fund Regulatory and Development Authority. Available at: https://www.pfrda.org.in/myauth/admin/showimg.cshtml?ID=433.

Pradhan, R. P. et al. (2015) 'Bond market development and economic growth: The G-20 experience', Tékhne, 13(1), pp. 51–65. doi: https://doi.org/10.1016/j.tekhne.2015.09.003.

Romei, V. and Reed, J. (2019) 'The Asian century is set to begin', Financial Times, 26 March. Available at: https://www.ft.com/content/520cb6f6-2958-11e9-a5ab-ff8ef2b976c7.

Sarr, A. and Lybek, T. (2002) Measuring Liquidity in Financial Markets. USA: INTERNATIONAL MONETARY FUND. doi: https://doi.org/10.5089/9781451875577.001.

SEBI (2019) Corporate Bonds - Statistics, Securities Exchange Board of India. Available at: https://www.sebi.gov.in/statistics/corporate-bonds.html.

Sharma, K. (2001) 'The Underlying Constraints on Corporate Bond Market Development in Southeast Asia', World Development, 29(8), pp. 1405–1419. doi: https://doi.org/10.1016/S0305-750X(01)00043-2.

Smaoui, H., Grandes, M. and Akindele, A. (2017) 'The Determinants of Bond Market Development: Further Evidence from Emerging and Developed Countries', Emerging Markets Review, 32, pp. 148–167. doi: https://doi.org/10.1016/j.ememar.2017.06.003.

Statista (2019) Leading stock exchanges in APAC 2018. Available at: https://www.statista.com/statistics/265236/domestic-market-capitalization-in-the-asia-pacific-region/ (Accessed: 20 January 2020).

Teplova, T. V. and Sokolova, T. V. (2018) 'Market Development Determinants for Corporate Bonds in National Currencies: Emerging Markets Review', Journal of East-West Business. Taylor & Francis, 24(1), pp. 50–80. doi: 10.1080/10669868.2017.1340387.

The World Bank (2006) 'Financial Sector Development Indicators', pp. 1–10. Available at: http://siteresources.worldbank.org/INTTOPACCFINSER/Resources/Bndind.pdf.

The World Bank (2019) Global Financial Development. Available at: https://databank.worldbank.org/source/global-financial-development#

2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)
Galgotias University, Greater Noida, UP, India. Oct 2-4, 2020

Melanoma Risk Prediction with respect to Modifiable Lifestyle Factors by Meta-Analysis Aided Machine Learning Technique

Avni Mehta
Department of Biotechnology
Delhi Technological University
Delhi, India
avnimehta28@gmail.com

Richa Nayak
Department of Biotechnology
Delhi Technological University
Delhi, India
ricn1996@gmail.com

Yasha Hasija
Department of Biotechnology
Delhi Technological University
Delhi, India
yashahasija06@gmail.com

***Abstract*—Melanoma is a highly prevalent dermatological disease and a life-threatening form of skin cancer. Lifestyle factors have been observed to influence melanoma risk. The relationship between increased BMI and melanoma incidence has been investigated using epidemiological, in our previous studies. In order to study relationships between melanoma and other modifiable lifestyle factors such as alcohol consumption, smoking, sunscreen application, and use of tanning devices, a meta-analysis aided by machine learning technique was carried out. PubMed database was searched carefully to sort out literature pertaining to the associations of these lifestyle factors with melanoma risk. Meta-analyses were carried out using a software called Review Manager 5.3. It gave risk ratios and 95% confidence intervals as results. Analysis of these data revealed weak positive relationship between alcohol consumption (OR = 1.46; 95% CI = 1.32 - 1.62) and use of tanning devices (OR = 1.36; 95% CI = 1.20 - 1.53). Negative associations were found between smoking (OR = 0.78; 95% CI = 0.67 – 0.92) and sunscreen application (OR = 0.5; 95% CI = 0.44-0.57). These results were matched, and association trends were confirmed with the rule induction results obtained by applying Naïve Bayes model to the data of each lifestyle factor. All the data from the five factors were pooled in together to create a master datasheet, and machine learning was performed on it to generate a predictive model for melanoma risk. The results were validated through a test split (ratio 0.7:0.3) and cross-validation as well. The accuracy was observed to be 70.23% and 70.35% + 0.79%. The functioning of the model was tested on an unlabeled dataset.**

Keywords—melanoma, machine learning, meta-analysis.

I. INTRODUCTION

Melanoma is a potentially fatal form of skin malignancy, and it is on the rise[1]. The incidence rates are increasing over the last decade, ranking 19th amongst the most common type of cancer worldwide[2], [3], more than that of other solid cancers[4]. According to the World Health Organization, 132,000 melanomas cases occur globally each year. An estimated 55,500 deaths are caused by cutaneous melanoma annually[5]. Risk of melanoma is associated with a multitude of factors like geography, ethnicity, age, sex, the incidence of Type 2 diabetes, Parkinson's disease, inflammatory bowel disease, exposure to UV radiation, genetics, and incidence of dysplastic nevus syndrome, and various lifestyle factors. Extensive research on melanoma and its risk factors have been carried out over the past decades, but the complex role played by modifiable lifestyle factors are largely unexplored. Several lifestyle factors that affect melanoma are obesity, alcohol consumption, smoking, sunscreen application, intake of carotenoids, selenium, coffee, etc. Meta-analysis is a comprehensive and systematic analysis of previous studies and pre-existing evidence present on a disease[6]. These kinds of studies can set a premise for further research by providing precise estimates of risk factors, treatment outcomes, and challenges. Since literature and data available from research and studies are often extensive, complex, and sometimes conflicting, it is important to perform a thorough analysis. Machine learning techniques, along with meta-analysis, can be used to perform such analysis in a precise and robust manner. The goal of this study is to build, as effectively as possible, a predictive machine learning model to determine the susceptibility of a person to develop melanoma with certain a lifestyle. Thus, some modifiable lifestyle factors were selected, such as obesity, the use of tanning devices, alcohol consumption, smoking, and sunscreen application.

II. OBJECTIVE

The relationship between the selected modifiable lifestyle factors and melanoma has been previously explored through meta-analyses. In order to confirm previous findings, we utilized recent epidemiological evidence to study these associations. The objective was to create a novel machine learning model to predict melanoma risk in people who follow a certain lifestyle, and findings from the meta-analyses were used to confirm these results. The disease melanoma was chosen because it is a dermatological disease with the lowest survival rate.

III. METHODOLOGY

A. Meta-analysis

1) Search Criteria: An extensive literature search was carried out through PubMed resources to evaluate the incidence of melanoma in persons with certain lifestyle habits concerning alcohol consumption [7]–[11], smoking[12]–[14], use of tanning devices[15]–[18] and sunscreen application[19]–[22]. Our search criteria included key terms like- "Melanoma" (e.g., "cutaneous melanoma," "skin neoplasm," "malignancy," "skin cancer," "risk factors"), and similarly for alcohol consumption, smoking, use of tanning devices and sunscreen application.

2) Selection of Study for Meta-analysis: Research papers with incomplete or inadequate data for calculating Hazard Ratio (HR) and 95% Confidence Interval (CI) were not included. Studies were evaluated to ensure that only

978-1-7281-5070-3/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:26:33 UTC from IEEE Xplore. Restrictions apply.

authentic studies were taken into consideration. Any study with duplicate or irrelevant data was excluded, and only studies published in English were taken into consideration. In studies with overlapping test populations, studies with a larger subset of cases and controls were filtered out for consideration. Studies not pertaining to melanoma and risk factors in consideration, in any way, were excluded. With a thorough examination of the studies, set inclusion or exclusion criteria were decided upon for facilitating the final decision.

3) Data Extraction: All the eligible studies were assessed, and the data retrieved were independently organized. For every study, baseline characteristics were determined. For example – Name of the first author, year of publication, affiliation, country, age range, sex, number of cases and controls, sources, and other adjustment factors[23]. In order to ensure that the data upon consideration is adequate for meta-analyses, the corresponding authors were contacted at least twice.

4) Statistical Analysis: Review Manager 5.3 was used to carry out statistical analyses. The odds ratios (ORs) were used as an indicator of the relationship between melanoma and the particular lifestyle factor. I^2 (a statistical measure that describes the percentage of variation owing to heterogeneity than chance[24]) value was used as a relative indicator for the measurement and quantification of heterogeneity of the candidate research papers. The method used to calculate the weighted odds ratios was the Mantel-Haenszel method[25], [26]. It is also known as the fixed-effect meta-analysis. $P > 0.1$ implied there is very little heterogeneity and the other way round. Heavy drinkers were classified as those who consume alcohol more than two times a week. Current and ex-smokers were classified as 'smoking' and those who have never smoked, as 'no smoking.' Based on the usage of tanning devices, people were classified into 'ever' or 'never' category. Those who apply sunscreen at least 3-4 times a week were placed under the category 'Sunscreen.'

B. Machine Learning-based Predictive Modelling

RapidMiner Studio 9.2, a data-science platform, was used to carry out machine learning on the datasets. It facilitates the hasslefree deployment of a predictive model by integrating data preparation and machine learning in a single workflow. Naïve Bayes classifier was used, which is one of the simplest Bayesian classifiers used in probabilistic machine learning models. Using this classifier algorithm, we will classify whether or not a person is susceptible to developing melanoma, given the lifestyle features. After designing of each process, the program is run to obtain results.

- For individual modeling for lifestyle factors, the following operators were used –

Retrieve → Set Role – 'Incidence' (Melanoma or no melanoma) was the label → Naïve Bayes → Rule Induction

Like in Bayes theorem, we can find the probability of melanoma occurring given the condition that one of the lifestyle features under consideration is adopted. The assumption being, all the lifestyle features under consideration are independent of each other.

- Pooling of data together

Once the results from individual modeling matched with the meta-analyses results, all the data was pooled in together to create a master datasheet of 36,310 entries.

- For test splits and validation, the following operators were used

Retrieve → Set Role → Split Data → Naïve Bayes → Apply Model → Measure Performance

- Cross-validation

It is essential to validate the predictive model generated using machine learning algorithms. The data was split in the ratio 0.7:0.3, i.e., 70% of the data was used as the training set, and the rest 30% was taken as a test set. The performance of the machine learning model is evaluated using the test set.

- Sample scoring

Now, with the machine learning model built, scoring, otherwise known as prediction is carried out on a new dataset in order to gain practical insights.

The model is now trained and ready for application for prediction and analysis.

IV. Results

A. Meta-analysis

1) Association between alcohol consumption and melanoma

Five studies were finalized to study the association between alcohol consumption and the risk of melanoma, as shown in Fig. 1. The final pooled estimate gave a significant result (OR = 1.46; 95% CI = 1.32 - 1.62). The heterogeneity results were also significant (I2 = 93%; P<0.00001).

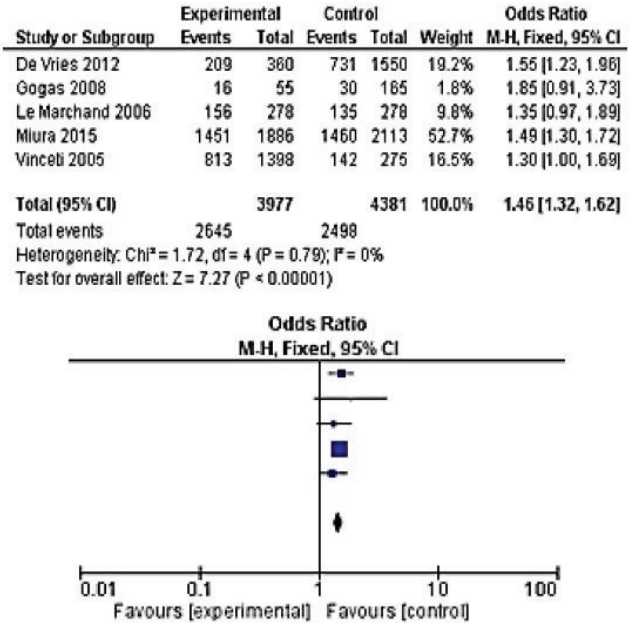


Fig. 1. Effect estimate of melanoma in people consuming alcohol, pooled together from various study and forest plot depicting odds ratio in 95% CI.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:26:33 UTC from IEEE Xplore. Restrictions apply.

2) Association between melanoma and the use of tanning devices

Four studies were finalized to study the association between tanning devices and melanoma risk as shown in Fig. 2. The final pooled estimate gave a significant result (OR = 1.36; 95% CI = 1.20 - 1.53). The heterogeneity results were also significant (I2 = 93%; P<0.00001).

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Bataille 2004	1120	1642	567	931	51.2%	1.38 [1.16, 1.63]
Clough-Gorr 2008	367	523	460	678	26.6%	1.11 [0.87, 1.43]
Cust 2010	137	604	84	479	16.1%	1.38 [1.02, 1.87]
Han 2006	42	182	87	712	6.1%	2.16 [1.43, 3.25]
Total (95% CI)		**2951**		**2800**	**100.0%**	**1.36 [1.20, 1.53]**
Total events	1666		1198			

Heterogeneity: Chi² = 7.33, df = 3 (P = 0.06); I² = 59%
Test for overall effect: Z = 4.93 (P < 0.00001)

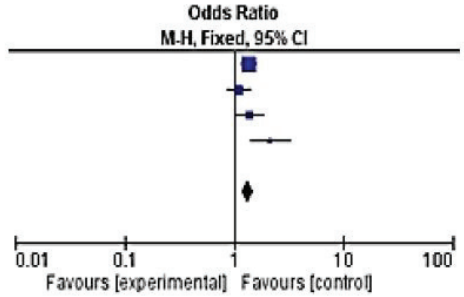


Fig. 2. Effect estimate of melanoma in people using tanning devices, pooled together from various study and forest plot depicting odds ratio in 95% CI.

3) Association between the use of smoking and melanoma

Seven studies were finalized to study the association between smoking and melanoma risk, as shown in Fig. 3. The final pooled estimate gave a significant result (OR = 0.78; 95% CI = 0.67 – 0.92). The heterogeneity results were also significant (I2 = 93%; P<0.00001).

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Antoniadis 2011	36	55	104	165	5.1%	1.11 [0.59, 2.11]
Batty 2009	404	815	68	122	16.9%	0.78 [0.53, 1.14]
De Hertog 2001	74	125	250	386	14.2%	0.79 [0.52, 1.19]
Fortes 2008	164	303	255	395	28.8%	0.65 [0.48, 0.88]
Gogas 2008	36	55	103	165	5.0%	1.14 [0.60, 2.16]
Kessides 2011	33	82	72	164	8.1%	0.86 [0.50, 1.47]
Le Marchand 2006	149	278	166	278	21.8%	0.78 [0.56, 1.09]
Total (95% CI)		**1713**		**1675**	**100.0%**	**0.78 [0.67, 0.92]**
Total events	896		1018			

Heterogeneity: Chi² = 4.08, df = 6 (P = 0.67); I² = 0%
Test for overall effect: Z = 3.03 (P = 0.002)

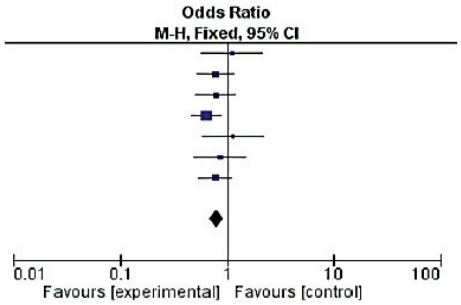


Fig. 3. Effect estimate of melanoma in people who smoke, pooled together from various study and forest plot depicting odds ratio in 95% CI.

4) Association between sunscreen application and melanoma

Four studies were finalized to study the association between the application of sunscreen and melanoma risk, as shown in Fig. 4. The final pooled estimate gave a significant result (OR = 0.5; 95% CI = 0.44 – 0.57). The heterogeneity results were also significant (I2 = 93%; P<0.00001).

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Green 2011	489	757	589	864	27.7%	0.85 [0.69, 1.05]
Klug 2010	460	717	743	945	32.7%	0.49 [0.39, 0.61]
Westerdahl 2000	127	658	123	410	17.4%	0.56 [0.42, 0.74]
Youl 2002	42	189	268	314	22.3%	0.05 [0.03, 0.08]
Total (95% CI)		**2321**		**2533**	**100.0%**	**0.50 [0.44, 0.57]**
Total events	1118		1723			

Heterogeneity: Chi² = 122.22, df = 3 (P < 0.00001); I² = 98%
Test for overall effect: Z = 10.93 (P < 0.00001)

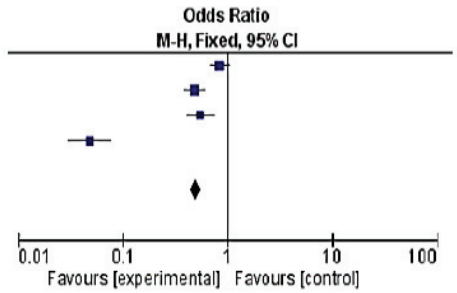


Fig. 4. Effect estimate of melanoma in people who use sunscreen, pooled together from various study and forest plot depicting odds ratio in 95% CI.

5) Exploration of heterogeneity

We evaluated heterogeneity by assessing the resulting forest plots visually. A Chi-square test was conducted. Non-overlapping CIs and P-values were generated. P < 0.1 implies that the result obtained is statistically significant, whereas I2 > 50% denotes sufficient heterogeneity.

6) Publication bias

Publication bias cannot be applied as each meta-analysis caters to less than ten studies.

B. Machine Learning-based Predictive Modelling.

1) Individual Modeling for Lifestyle Factors

Through Rule Induction and Naïve Bayes, results showed a positive association between alcohol consumption and the use of tanning devices. Negative associations were found between smoking and sunscreen application. These results were coherent with the trends observed in the meta-analyses.

2) Test Splits and Validation

Accuracy of 70.23% and a precision of 70.29% was obtained.

3) Cross-Validation

An accuracy of 70.35% + 0.79%. was obtained.

4) Sample Scoring

An unlabeled dataset was used for further validation.

V. Discussion

The novelty of this study lies in the fact that it uses a two-layered system for analyses, i.e., meta-analyses followed by machine learning-based predictive modeling. Each of the two methods has its own individual criterion for

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:26:33 UTC from IEEE Xplore. Restrictions apply.

further validation. This affirms the authenticity of the results obtained since we proceeded with the second method only when proper results were obtained in the first method. For instance, there are modifiable lifestyle factors beyond the ones used in this study. However, since significant results were not obtained during the meta-analyses, these factors were not considered credible and were hence, excluded. As more literature is published, more and more factors can be incorporated, and the accuracy of this model will go up. The current model accuracy of ~70% can be considered to be sufficient. It would not have been worthy of consideration if the variables were smoking and lung cancer or anxiety levels and Parkinson's disease. Since the variables in question (melanoma VS factors like alcohol consumption, smoking, sunscreen application, and use of tanning devices) are not known to have direct links or causal relationships so far, this model is of importance to the scientific world.

Possible limitations of our meta-analyses can be attributed to shortcomings of individual studies and confounding bias. Exclusion criteria for filtering out studies in languages other than English, and articles only available through PubMed, limited variability in selected studies. It can be said that meta-analyses suffer from certain methodical limitations, such as information bias or selection bias. However, since this analysis was one of its kind and was based on high-quality studies, it opens up new avenues for research in the area.

VI. Conclusion

Lifestyle factors and the environment has a significant role to play in many complex diseases. However, the determination of the degree of its effect is challenging, so much so that it is often not taken into consideration for disease prognosis studies. With the current surge in evidence-based research, it is possible to perform such an analysis where we factor in lifestyle and environmental factors while studying diseases.

The predictive model developed by machine learning with the help of meta-analyses takes into account the lifestyle habits of a person in terms of obesity, smoking, alcohol consumption, sunscreen application, and use of tanning devices. However, since this is the first study to evaluate such a relationship, further confirmatory tests based on more literature should be carried out in order to get better results.

References

[1] A. Ingraffea, "Melanoma," Facial Plastic Surgery Clinics of North America, vol. 21, no. 1. Facial Plast Surg Clin North Am, pp. 33–42, Feb-2013.

[2] J. Ferlay et al., "Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods," International Journal of Cancer, vol. 144, no. 8. Wiley-Liss Inc., pp. 1941–1953, 15-Apr-2019.

[3] "Skin cancer statistics | World Cancer Research Fund." [Online]. Available: https://www.wcrf.org/dietandcancer/cancer-trends/skin-cancer-statistics. [Accessed: 29-Feb-2020].

[4] B. Owens, "Melanoma," Nature, vol. 515, no. 7527. Nature Publishing Group, p. S109, 20-Nov-2014.

[5] D. Schadendorf et al., "Melanoma," The Lancet. 2018.

[6] Meta-Analysis in Medical Research. 2005.

[7] K. Miura et al., "Alcohol consumption and risk of melanoma among women: pooled analysis of eight case-control studies," Arch. Dermatol. Res., 2015.

[8] E. De Vries et al., "Known and potential new risk factors for skin cancer in European populations: A multicentre case-control study," Br. J. Dermatol., 2012.

[9] H. Gogas et al., "Melanoma risk in association with serum leptin levels and lifestyle parameters: A case-control study," Ann. Oncol., 2008.

[10] L. Le Marchand et al., "Sun exposure, diet, and melanoma in Hawaii Caucasians," Am. J. Epidemiol., 2006.

[11] M. Vinceti et al., "A population-based case-control study of diet and melanoma risk in northern Italy," Public Health Nutr., 2005.

[12] G. D. Batty, M. Kivimaki, L. Gray, G. D. Smith, M. G. Marmot, and M. J. Shipley, "Cigarette smoking and site-specific cancer mortality: Testing uncertain associations using extended follow-up of the original Whitehall study," Ann. Oncol., 2008.

[13] S. A. E. De Hertog et al., "Relation between smoking and skin cancer," J. Clin. Oncol., 2001.

[14] M. C. Kessides, L. Wheless, J. Hoffman-Bolton, S. Clipp, R. M. Alani, and A. J. Alberg, "Cigarette smoking and malignant melanoma: A case-control study," J. Am. Acad. Dermatol., 2011.

[15] R. P. Gallagher, J. J. Spinelli, and T. K. Lee, "Tanning beds, sunlamps, and risk of cutaneous malignant melanoma," Cancer Epidemiology Biomarkers and Prevention. 2005.

[16] D. K. Woo and M. J. Eide, "Tanning beds, skin cancer, and vitamin D: An examination of the scientific evidence and public health implications," Dermatologic Therapy. 2010.

[17] W. Ting, K. Schultz, N. N. Cac, M. Peterson, and H. W. Walling, "Tanning bed exposure increases the risk of malignant melanoma," Int. J. Dermatol., 2007.

[18] K. M. Clough-Gorr, L. Titus-Ernstoff, A. E. Perry, S. K. Spencer, and M. S. Ernstoff, "Exposure to sunlamps, tanning beds, and melanoma risk," Cancer Causes Control, 2008.

[19] H. L. P. Klug et al., "Sunscreen prevention of melanoma in man and mouse," Pigment Cell and Melanoma Research. 2010.

[20] A. C. Green, G. M. Williams, V. Logan, and G. M. Strutton, "Reduced melanoma after regular sunscreen use: Randomized trial follow-up," J. Clin. Oncol., 2011.

[21] M. B. Planta, "Sunscreen and melanoma: Is our prevention message correct?," J. Am. Board Fam. Med., 2011.

[22] D. A. Lazovich, R. I. Vogel, M. Berwick, M. A. Weinstock, E. M. Warshaw, and K. E. Anderson, "Melanoma risk in relation to use of sunscreen or other sun protection methods," Cancer Epidemiol. Biomarkers Prev., 2011.

[23] A. Mehta and Y. Hasija, "Meta-analyses of associations of obesity with risk of melanoma and acne based on case-control studies," Biomed. Dermatology, 2019.

[24] J. P. T. Higgins and S. G. Thompson, "Quantifying heterogeneity in a meta-analysis," Stat. Med., 2002.

[25] N. Hamajima, K. Matsuo, and H. Yuasa, "Meta-analysis," Biotherapy, 2000.

[26] A. Mannocci, "The Mantel-Haenszel procedure. 50 years of the statistical method for confounders control," Ital. J. Public Health, 2009.

[27] Milan Kumar, V. M. Shenbagaraman and Ankush Ghosh, "Innovations in Electrical and Electronic Engineering" Book Chapter, Springer. [ISBN 978-981-15-4691-4, Favorskaya et al (Eds.): Innovations in Electrical...]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:26:33 UTC from IEEE Xplore. Restrictions apply.

Wireless Personal Communications
https://doi.org/10.1007/s11277-020-07879-x

mTrust: Call Behavioral Trust Predictive Analytics Using Unsupervised Learning in Mobile Cloud Computing

Arka Bhowmik[1] · **Debashis De**[2]

Accepted: 29 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
Filtering spam voice calls are a still major challenge in today's technology contrary to SMS or email-based spamming. A numerical measure of the trust between users can help us filter calls based on relevance. Given the abundance of user-generated information available from the huge number of online devices, we can harness the power of this data to develop software adapting to user behavior. Existing research works for trust computation face various challenges when it comes to global applicability and understandability of trust values. Our investigation includes detailed surveillance of user call patterns based on the call data available from mobile devices and proposes a novel approach to filter calls that are of higher relevance to users based on their call-trust values. Our implementation realizes the diversity in call patterns of different people due to varying usage and uses classification and clustering algorithms to generate personalized, accurate numerical, and categorical trust values for every user. Categorical trust makes it easier to apply and understand trust ratings on a global scale. The implementation also incorporates a cloud facility to crowdsource trust values from multiple users, in a single database to generate the global trust of a user which can be used for spam filtering on a global scale. A software named "mTrust" is developed in this work for the future generation of a trustworthy mobile cloud network.

Keywords Trust · Unsupervised learning · K-Means · Mobile cloud computing · Android application · Human behavior

✉ Debashis De
dr.debashis.de@gmail.com

Arka Bhowmik
arka.bhowmik95@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, New Delhi, Main Bawana Road, New Delhi 110042, India

[2] Centre of Mobile Cloud Computing, Department of Computer Science and Engineering, Maulana Abul Kalam Azad University of Technology, B.F.-142, Salt Lake, Sector-1, Kolkata, West Bengal 700064, India

Published online: 22 November 2020

1 Introduction

Trust is a measure of reliability and strength of a relationship. Although it is challenging to perfectly measure a person's trust in another person, with a numerical value, a user's behavior towards another person may provide us an inference on a value of trust between two individuals. The trust value may be calculated and defined within a given fixed range like 0–1; where 0 indicates low trust and 1. The process of computation and application of trust can be redefined for different scenarios of usage, for example, call-based trust in mobile networks, social-trust in microblogging websites, and document-trust [1] for determining document content relevance. However, all applications serve one common purpose, which is to determine the genuineness of the subject. Recent research works have found various other applications of trust values, insecurity [2, 3], recommender systems [4, 5], and social influencer identification. Numerically determining the trustworthiness of users can be quite a challenging task due to the diversified usage patterns of devices among different users. The two main steps towards the calculation of trust involve: firstly, to identify the important features of the scenario that can be associated with a value of trust and secondly, to devise an algorithm that can equate the features into generating an accurate value for trust. However, the feature-set alone cannot always provide us with a convincing value to determine trust, as it is also dependent on a user's behavior which can sometimes be hard to decipher. It is hence necessary to consider both behavioral and numerical parameters while determining the trust between two people in a network. A trust value thus generated may be convincing from the contributing user's perspective. However, this value also has to be normalized concerning all other users in the network, so that it can be used globally. Considering all these factors, on the successful calculation of a global trust value, from the contribution from multiple subscribed users, against a particular subscriber, we can classify that person's trustworthiness on a global scale. Dynamic Trust is a property of the trust to adapt to a changing environment, be it re-computation of trust in a single device over varying intervals of time or updating global trust in a server with varying user interactions between connected people, over time. Depending on the type of trust under consideration, we have defined various methods to calculate them.

1.1 Numerical Trust

Numerical Trust can be calculated based on several valued parameters calculated from available data that may be related to calls, or message interactions, or other social activities. The procedure for the calculation can be based on IF–THEN–ELSE statements and mathematical equations. However, machine learning methods cannot determine fuzzy values for numerical trust.

1.2 Categorical Trust

Categorical Trust requires generating values of trust corresponding to various classes viz: low, medium, high. Values thus generated are easier to understand. They can be used in Machine Learning and AI methods to train a network to predict the trust of an unknown record. Some methods which can be used for predicting categorical trust are.

1.2.1 Fuzzy Logic

Fuzzy Logic, instead of making hard assignments of the category of trust, makes probabilistic assignments to the same. This model can be beneficial for discriminating ambiguous records from the database where we cannot ascertain the category of trust to a record, with high confidence. Fuzzy [6] Logic expressions are similar to that of Boolean Algebra.

1.2.2 Decision Trees

Decision Trees [7] are sub-optimal trees where at every level, we have a split based on a particular feature, deciding the predictions for the next level. At every level, the entropy of the network decreases, which ensures that the decision of the tree is enhancing in every step. Having multiple decision splits based on call attributes can help us categorize the record being spam or not.

1.2.3 Classifiers

If we have sufficient training data, we can analyze and extract relevant features from the data and train a classifier like a Perceptron model [8], SVM [9], or KNN [10] to classify data records based on its features. Such a network will be able to continuously improve itself with more data being gradually fed from a user network.

1.3 Motivation and Contributions

Our implementation consists of an android application that uses call logs to extract data in a given interval of time to calculate trust and then updates this data on the global server. The global trust can be calculated aggregating the values contributed towards that person from other people who have contacted him. Once the records have been updated, they can be accessed from local devices. The whole architecture is described in Fig. 1. Our work contributes towards the following based on two different methods:

1. A new approach in determining trust, which is both easily comprehensible and applicable in a global environment, using clustering and classification. The records per user will be stored in a cloud database facility called Trust Database for preserving a user's trust on a global level.
2. A solution for the continued enhancement of our model using a user feedback system based on the crowd-sourced rating of users based on the level of trust.
3. Generating personalized classifiers per device, for personalized trust.
4. Redefining behavioral parameters for a dynamic approach towards trust computation.

The rest of the paper has been organized as follows. Section 2 contains the research work that is related to our project and inspired the implementation of this work. Section 3 describes our proposed model with two different implementations. The first method involves an improved implementation of Numerical trust computation while the second method describes a novel approach towards categorical Trust calculation, to deal with the shortcomings of numerical trust-based approaches. Section 4 describes the implementation process of both the proposed methods, along with a detailed description of the dataset used to train the model used in the second method. Section 5 contains a detailed analysis of the

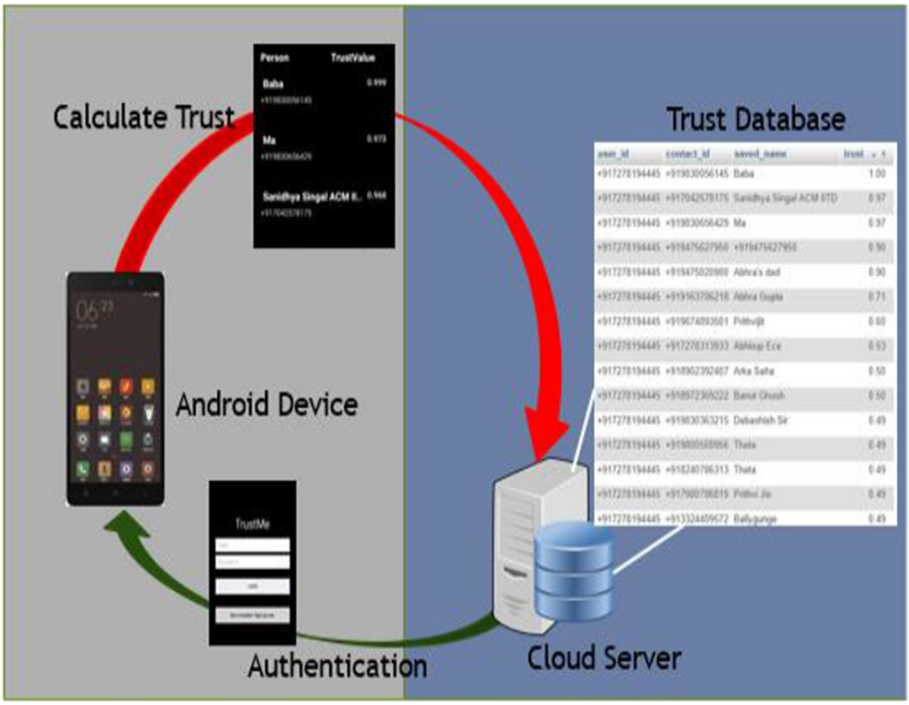


Fig. 1 The architecture of mTrust software design in mobile cloud computing

data associated with the research, results of the proposed work, and application screenshots. This section is followed by future work proposals and the conclusion of this research.

2 Related Work

Global- Social Trust ensures a more reliable solution towards the calculation of trust and various research has been conducted regarding this. Shuhong Chen et al. devised a model [11] for computing trust on a global scale on multiple clouds by the aggregating local trust from multiple devices. The transitive relationships are taken into account so that the global trust is not reflected by a single person's experience of another user. In a previous research D, De, and B Dey have devised a model called TBFN [12] which computes global trust based on both SMS and Call parameters. The model emphasizes on the level Hierarchy of trust parameters and also multiple methods for computing global trust depending on the number of short-paths created to a person in a social network. Apart from call-based trust, the trust may be also directly assigned by users, in a global environment. Liu Xin et al. describes a Dynamic Trust Algorithm for Social Networks [13] where a social graph is created from the user assigned trust and connected nodes influence the trust of their direct neighbors. Their model describes a dynamic approach towards adjusting global trust when subject to any changes. Trust predictions can also be based on the trust propagation dynamics within a network based on the global and direct spectrum. Hyun-Kyo in [14] describes how their IDT (Indirect, Direct, and Transposed trust) model outperformed

other models using trust propagation dynamics in predicting trust. Top k social connections with adjusted weightage can contribute towards calculation of social trust which has been proposed in [15] as a measure of similarity between people to design a recommender system, based on Sina microblog data. Adali's paper [16] uses twitter messages to determine trust by emphasizing in Length, frequency, and balancing of twitter conversations. Mucheol Kim and Sang Oh Park's paper [17] titled "Trust management on user behavioral patterns for a mobile cloud computing" serves as a big inspiration for our paper in selecting the feature-set for determining trust value. They propose a social-cloud based trust model using frequency, intimacy, and recency of call as attributes to generate a fuzzy value ($0 \leq \text{trust} < 1$) that will determine the trustworthiness of other users. However, using this method, there are a few challenges both in globally scaled and normalized trust values as well as applicability in a dynamic environment. For proper weight assignment to decision attributes a recommendation trust solution has been suggested by Bin Zhao in his work [18]. Before assigning global trust, it is extremely important to get the local trust calculation accurate and understandable. To tackle this problem and enhance the generation of more accurate trust values, we propose a novel implementation using non-linear algorithms like Multi-Layer-Perceptron, to generate trust values in a set of ranged categories which are both, easier to understand and evaluate. Generated labels can be further enhanced by getting them relabelled to their expected values by users. This type of crowd-sourcing [19] labels enhance the quality of future predictions in the network.

3 Proposed Model of mTrust

In this section, we discuss the 2 approaches for trust calculation with its working principle. The mathematical modeling is also described.

We choose parameters from papers [11, 17] to computer trust in our work. However, in this paper, some significant changes have been made in some of the computation methods of the reference papers, notably the calculation of recency and normalization of the features. The trust model mainly comprises of 3 features of a person's calls, namely: frequency, intimacy, and recency.

Frequency The number of calls made to a person gives us significant information regarding the importance of a person in someone's life. The number of outgoing calls made to user B from user A emphasizes how much user A trusts user B. the first attribute is hence computed by the number of outgoing-calls users A makes to user B, relative to the total number of outgoing calls he has made in that period. If C_{AB} is the number of outgoing calls from A to B and C_{AK} is the number of calls from A to K where K is a set of n total users:

$$Frequency = \frac{C_{AB}}{\sum_{k=1}^{n} C_{AK}} \tag{1}$$

Intimacy A person's intimacy with another person is the measure of the closeness of that person to the user. The duration of a call determines the intimacy of the relationship between two people. People tend to have higher duration calls with people who are closer to them. The intimacy can rely on both incoming and outgoing calls, the weightage of their contributions being based on whether the user is a sending-oriented or reception-oriented person. To determine the call orientation of a person, we use the attribute "tendency". Tendency finds its significance in ascertaining the correct weightage that is to be given to

outgoing and incoming calls for every unique person since the frequency of the types of calls varies hugely from person to person. If D_{AK} denotes call duration of outgoing calls from A to K and D_{KA} is the duration of incoming calls from K to A, and there are a total n number of users:

$$Tendency = \frac{\sum_{k=1}^{n} D_{AK}}{\sum_{k=1}^{n} D_{AK} + \sum_{k=1}^{n} D_{KA}} \tag{2}$$

S_{AB} is a relative measure of sending oriented duration of A to B.

$$S_{AB} = \frac{\sum_{k=1}^{n} D_{AB}}{\sum_{k=1}^{n} D_{AK} + \sum_{k=1}^{n} D_{KA}} \tag{3}$$

R_{BA} is a relative measure of the reception-oriented duration of B concerning A.

$$R_{BA} = \frac{\sum_{k=1}^{n} D_{BA}}{\sum_{k=1}^{n} D_{AK} + \sum_{k=1}^{n} D_{KA}} \tag{4}$$

Finally, Intimacy is calculated as

$$Intimacy = Tendency * S_{AB} + (1 - Tendency) * R_{BA} \tag{5}$$

Recency Recency of a person is the measure of how recently person A has interacted with person B. For simplified and more accurate results, we have normalized the dates based on the earliest call ever made on the device. If T_{AB} denotes the timestamp of a call (in milliseconds) made between A and B

$$Recency = \frac{T_{AB} - Min_{k=1}^{n}(T_{AK})}{Max_{k=1}^{n}(T_{AK}) - Min_{k=1}^{n}(T_{AK})} \tag{6}$$

3.1 METHOD I: Numerical Trust

In this method, we try to have a linear association between the features to determine the trust. To determine the weightage that should be given to each of the 3 features, we have the 3 weighing factors alpha, beta, and gamma, which are constants for a particular instance of the program. These are behavioral constants that vary from person to person and also with time, depending a person's current interaction. The sum of the weights of alpha–beta and gamma will add up to 1. However, the individual weights are calculated separately. For a given period:

α is the ratio of the outgoing calls to total calls:

$$\alpha' = \frac{Total\ number\ of\ outgoing\ calls}{Total\ number\ of\ calls} \tag{7}$$

β is a measure of the time the user invests in calls. Statistics show that an average person makes calls of duration ranging from 50 to 80 s per call and calls having greater or lesser duration are few and can be termed as outliers. If the distribution of the call duration is assumed to follow a Gaussian distribution approximately, β is calculated as:

$$\beta' = \frac{\mu(mean\ duration)}{NMD\ (normalized\ maximum\ duration)} \tag{8}$$

$$NMD = \mu + 3 * \sigma \tag{9}$$

where σ is the standard deviation of call durations.

γ is a measure of how active has the person been concerning phone calls, recently.

$$\gamma' = \frac{\frac{1}{N}\sum_{k=1}^{N}(T_{AK} - Min_{k=1}^{n}(T_{AK}))}{Max_{k=1}^{n}(T_{AK}) - Min_{k=1}^{n}(T_{AK})} \tag{10}$$

$$Now\ \alpha + \beta + \gamma = 1\ and\ 0 \leq \alpha,\ \beta, \gamma \leq\ 1 \tag{11}$$

After calculating α′, β' and γ′ separately, we take their weighted values as the actual values for α, β, and γ. Finally, trust is calculated as follows:

$$Trust_{AB} = \alpha * Frequency_{A,B} + \beta * Intimacy_{A,B} + \gamma * Recency_{A,B} \tag{12}$$

3.2 METHOD II: Categorical Trust

This method adopts a non-linear association amongst the features for determining categorical trust. So, we do not need α, β, *and* γ here. To handle non-linearity, we have to change our model from fuzzy to categorical, so that a classifier can be trained to determine the class of trust, a user belongs to. To train a model, we needed to label our data beforehand, for which we used a K-Means unsupervised clustering algorithm. K-Means algorithm plots raw data in a hypothetical k dimensional space to determine groups of points which are located more closely and form a cluster with those set of points. The number of clusters formed is equal to the number of classes we require. In every iteration of cluster formation, the cluster heads keep on shifting till the setup is stable and no datapoint changes its cluster association anymore. The cluster labels are then used as training labels during training, which is performed in a Multi-Layer Perceptron model. The perceptron is capable of automatically determining the weightage division among features to properly assign labels to our data. It is beyond the scope of hard assignments or any inductive assignments to achieve this. The class having the highest score in the perceptron output will be the proper predicted class of the record. The overall process is summarized in Fig. 2.

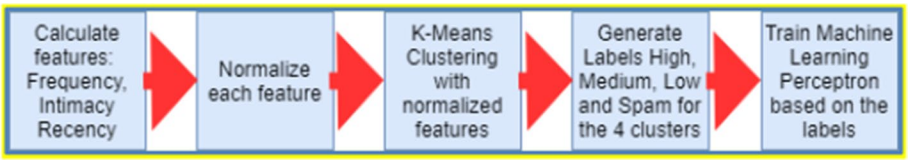


Fig. 2 Clustering and training process for categorical trust

4 Implementation of mTrust

The implementation has been carried out in an android application integrated with an online server. The server uses MySQL database with InnoDB Engine and PHP scripting as back-end. The android application uses the TensorFlow Android library for trust prediction and We extract the call features from the android call log database. The application also has feedback support for users to review and send the data back to the global server so that the prediction model performance can be enhanced. The application has been primarily tested on a Redmi 3S mobile device with 3 GB RAM, running on Android 6.0 Marshmallow.

4.1 Cloud Setup for mTrust

To make the most out of a person's collected data and also to validate the user's authenticity, an authentication model has also been introduced in the application which requires the user to login with valid credentials to access online features of the application. The online features of the application include the profile of the user and also access to the cloud data which holds records of people with their names, numbers, and trust. The overall database schema of the global server is described in Fig. 3.

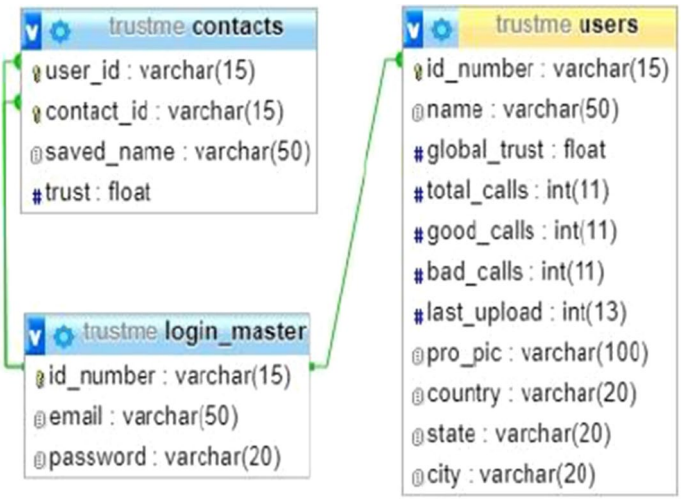


Fig. 3 Schema of the global server

4.2 Algorithm for mTrust

Input: Call log, Name, Number.
Output: Set of the trust value

1. **Start**
2. **If** a user is already logged in
   a. Jump to
3. **Else If** User is already registered
   a. Input credentials username and password
   b. Credentials sent to a server
   c. The server checks record in a database and authenticates a user
   d. **If** credentials wrong
      i. **Jump to** 1
   e. **Else**
      i. Jump to 6
   f. **End If**
4. **Else**
   a. User has to register with valid credentials into the app
   b. Details sent to the server
   c. **Jump to**
5. **End If**
6. **For** every call-data from call log:
   a. Get user number, call duration, type, date timestamp
   b. Accumulate all details in a Person Object, hashed to the Contact Number of that user
   c. Update most recent and least recent calls
   d. Update maximum duration, minimum duration calls
   End For
7. **Calculate** basic trust behavioral features
8. **Calculate** call statistics like mean, SD of call duration and frequency
9. **Calculate** frequency, intimacy, and recency
10. **Calculate Trust** based on user preferred method (Categorical or Numerical)
11. **Store** all values in the local database
12. **Display** the data to the user
13. **Upload** the generated data to the cloud
14. **On receiving** new data in the cloud
    a. **For** $Person_k$ in List of Persons updated by current User:
       i. **Update or Create** a record of $Person_k$ in Server Database
       ii. **If** Categorical Trust is used
           Get Votes for every Trust category for $Person_k$ from database
           Update $Person_k$ Trust based on Maximum Votes
       iii. **Else If** Numerical Trust is used
           Update $Person_k$ Trust as an average of Mutual trust from all connected users
       iv. End If
    b. **End For**
15. **Stop**

The overall process is shown in the Activity diagram in Fig. 4. Figure 5 demonstrates the salient features of the designed android application.

The internal algorithm for calculation has some alterations for Method 1 and method 2, which are mutually exclusive to each other. These changes are stated in Sects. 4.3 and 4.4.

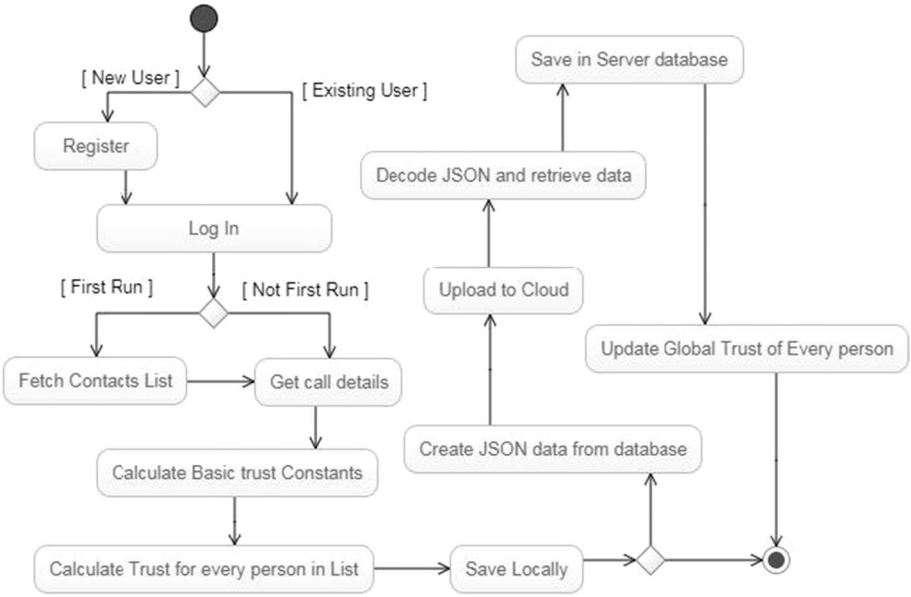


Fig. 4 Activity diagram of the trust evaluation process

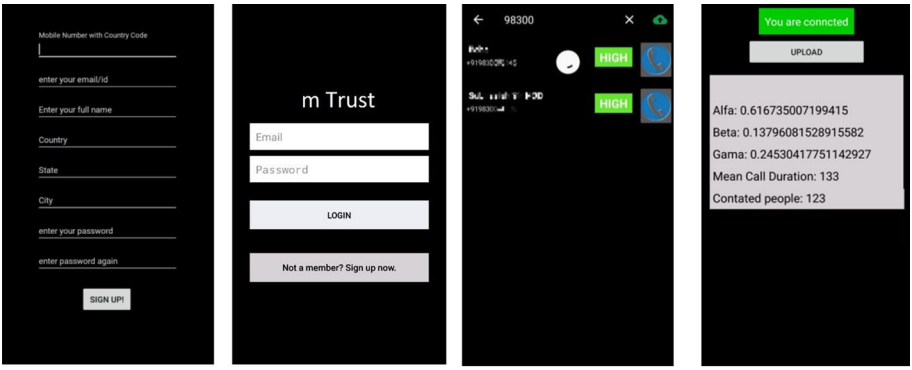


Fig. 5 Salient features of the developed application (**a**). Register screen, (**b**). Login screen, (**c**). Unidentified number detection from the global database, (**d**). User upload generated data to the cloud

4.3 Implementation for Method I

To ensure that the values are updated properly, on the first run, we consider calls not older than 100 days to calculate the values. On later trials, only the calls made on dates later than the last updated date are considered for evaluation and the corresponding trust values are updated. The changes in the trust values of people who have not participated recently are unaltered in this case. However, if the period of their dormancy is considerably significant, their trust values need to be decreased.

Fine-Tuning For getting a better approximation for the ratio γ' we need to reduce the range from milliseconds to days (of both numerator and denominator). Higher ranges of γ' (> 0.9) indicate that the user has been active recently and all records are more or less recent since the last updated date. Hence the calls should be differentiated from each-other mainly based on frequency and recency. Thus, the value for γ should be decreased to adjust the weightage of α and β. Again, smaller values of γ' (< 0.3) indicate that the user has not been active recently and again α and β should be given more weightage than γ. The ideal value for γ' should be in the range of 0.45–0.75 to give balanced weightage to recency while determining Trust. The adjustments are made accordingly.

4.4 Implementation for Method II

In this section, we explore the possibility of non-linear relationships between the chosen attributes of recency, frequency, and intimacy for trust calculation and describe the process of generating a categorical trust value instead of fuzzy or decimal trust, by training a machine learning model with unsupervised learning based on a real-life dataset of call and SMS records.

4.4.1 Dataset and Pre-processing

We have used a database of 65,000 detailed call and SMS records containing data of around 6500 unique subscribed numbers. To filter the raw data based on our requirements we had to perform rigorous preprocessing using the Pandas library in Python. Only voice calls were taken into consideration with the following attributes in focus: calling number, called the number, direction, duration, date of the call. The database contains around 39,000 voice calls of which 6926 are roaming calls. In the next step, we formulated two other tables from the mother table. Table 1 consisted of user data containing the total outward and inward durations, total outgoing and incoming calls, and his tendency of being a call sending oriented person. Table 2 consisted of the mutual call records between two persons, based on total outgoing and incoming calls from the other person, in-call and out-call

Table 1 Comparison of contributions of proposed work with related work

Features: other papers	Dynamic trust value detection	Customized behavioral constants	Personalized trust classifier per device	Global trust determination on Cloud Database	Feedback based trust
Our implementation	✓	✓	✓	✓	✓
Trust management on user behavioral patterns for mobile cloud computing [17]	✓	×	×	✓	×
Trust-based VoIP Spam detection based on call duration and human relationships [20]	✓	×	×	✓	×
A trust model using implicit call behavioral graph for mobile cloud computing [11]	✓	×	×	✓	×
A dynamic trust conference algorithm for social network [13]	✓	×	×	✓	✓

Table 2 Description of parameters used in equations

Parameter	Description
C_{AB}	The number of outgoing calls from person A to person B
D_{AB}	The duration of outgoing calls from person A to person B
S_{AB}	Sender Oriented duration relative to total call duration
R_{BA}	Reception Oriented duration relative to total call duration
T_{AB}	Timestamp of a call from A to B
μ	Mean Duration of a call
σ	Standard Deviation of call duration
α	Normalized Call Behavioral constant for call frequency-based feature
β	Normalized Call Behavioral constant for call intimacy-based feature
γ	Normalized Call Behavioral constant for call recency-based feature

durations, and the timestamp of the most recent call made between the two persons. While pre-processing it was important to note a few points:

- Zero or One duration calls (missed-calls) have been ignored while calculating average call duration and the total duration of incoming and outgoing calls since they can cause ambiguity in determining a relative measure of the calls which were picked up by the users.
- However, zero duration calls, if of incoming type, have been considered for total call frequency count between the two parties, as people do not intend to pick up calls from numbers which they do not trust.
- The call dates have been converted into their timestamp values for simplified comparison.

4.4.2 Labelling Data

After feature (frequency, intimacy, recency) extraction, we had around 13,500 entries for mutual interactions between different people. To label the trust categories for the data, we have used K-Means [21] clustering algorithm from the sci-kit learn library in python. From the diagram in Fig. 6, we have 4135 low trusts, 3235 medium trusts, 3150 high trust, and 2886 spam calls. Here low trust denotes people who have been in less contact. The low trust (Red) cluster is quite diverse while the other clusters are quite tight. Before clustering, for normalization of data, we have used the standard score instead of feature scaling. Standard scores are suitable when the data follow a normal distribution. In our case feature scaling had no impact in normalizing, as the maximum and minimum values for any feature were 0 and 1 only.

Standard score	Feature scaling
$X' = \frac{X-\mu}{\sigma}$	$X' = \frac{X-X_{min}}{X_{max}-X_{min}}$

The cluster labels thus generated are used to label trust categories in the dataset. Since we have 3 features, it is possible to have an understandable visual representation of the clusters formed in a 3-dimensional plane, which is demonstrated in Fig. 6.

4.4.3 Training of Multilayer Perceptron

Based on labels generated we trained a Multi-Layer Perceptron (MLP) [22] model in Keras. A simple MLP was chosen so that the model keeping in mind the computational limitations in a mobile device. After ensuring almost immaculate accuracies by the model, we transferred our model into android's TensorFlow integration to predict values in our android application. Figure 7 shows the exact depiction of our MLP.

5 Results and Discussions

Studying the records of the dataset, we find that there is great diversity in call patterns of users Following histograms show the various call patterns of people from the database. Figure 8a shows the user's participation level in outgoing calls and Fig. 8b visualizes the duration of calls in the database. Here, we have filtered outlying records as the main focus was on the average user behavior.

For any training process, the feature-set attributes need to have minimal to almost no correlation among each-other. We plot the frequency, recency, and intimacy parameters amongst each other in Fig. 9, to find that there is no possible pattern amongst them. This ensures our choice of these 3 features are un-correlated and hence serve as a good choice for training.

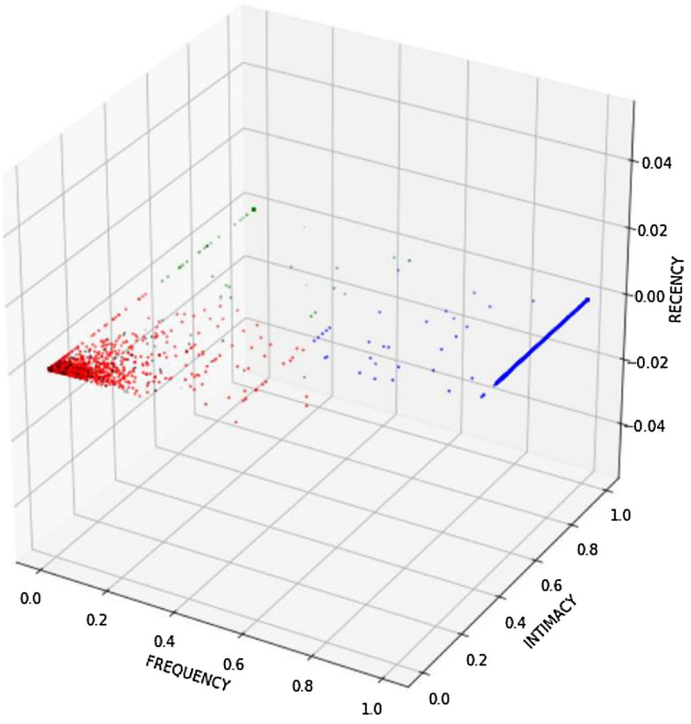


Fig. 6 Clustering along 3D axes with 4 clusters [Black: Spam, Red: Low trust, Green: Medium trust, Blue: High Trust]. (Color figure online)

The mean and standard deviation of the computed parameters from the dataset are represented in Table 3. Statistically better readings are available if values are calculated over a longer course of time

Using method 1, the numerical local trust and global trust are shown in Fig. 10a, b respectively. The names and numbers have been obfuscated to preserve the privacy of the users. The calculated values of α, β, and γ for an arbitrary user are shown in Fig. 10c. For most cases, a trust value of 0.5 is stated as high, and 0.1 is stated to be low. Values can change with time.

For Method 2, while training the perceptron on our cluster-labeled dataset with 13,500 entries, we got a validation set prediction accuracy of 99.88%. Thus, ensuring the sharp accuracy of our model, the model was exported to the android platform using TensorFlow. The application performance is quite satisfactory.

The generated Categorical Trust values by Method 2, are shown in Fig. 11a. We propose an approach to build truly personalized prediction models, trained with user-specific feedback data on top of our pre-trained model, using transfer learning. This is presumed to predict personal trust more accurately, on a user level. Hence, we also have feedback support for the application where users can re-label the predicted data and send it back to the servers. The feedback interface is shown in Fig. 11b. The labels High, Medium Low, and Spam define the strength of the trust of the user on a personal level. The star ratings, on a scale of 1–4 (1 indicating spam and 4 indicating High Trust) indicate the rating the user manually uploads for the corresponding contact in our feedback interface, for enhancing the trained perceptron model which is generated on his device on a personal level for better trust rating predictions in the future.

6 Future Scope

Some of the future research challenges for Mobile cloud trust management are discussed as open research problems.

Challenge 1: Better predictions with more data

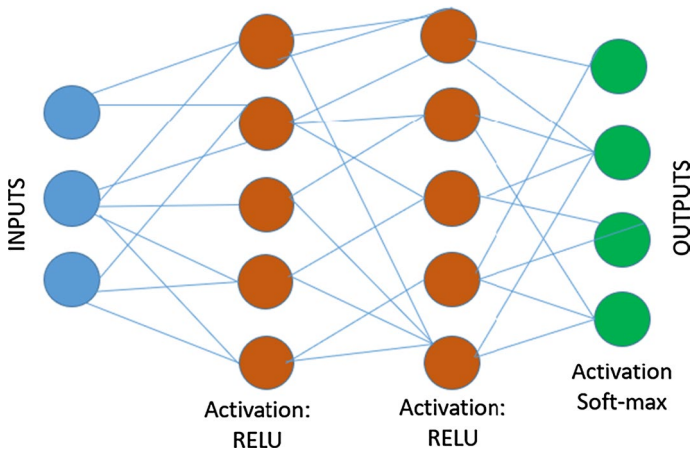


Fig. 7 The MLP model for trust prediction

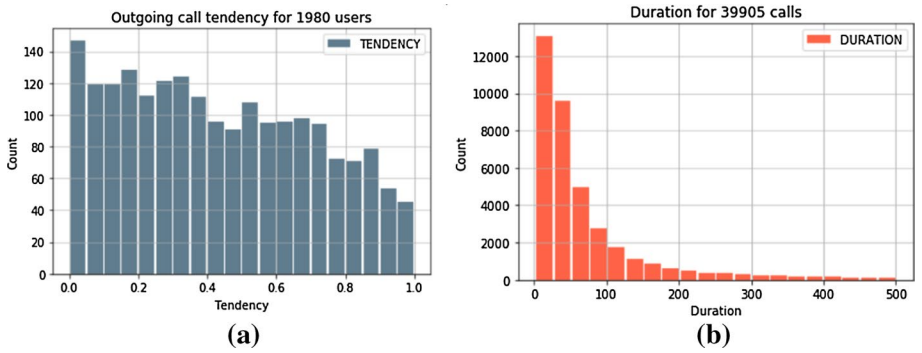


Fig. 8 **a** Histogram showing outgoing call tendency of users, **b** histogram showing duration (in a s) of the calls in our dataset

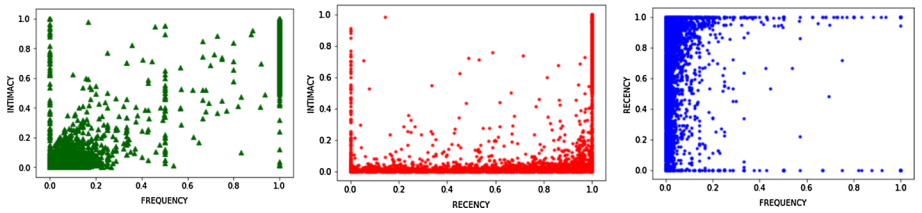


Fig. 9 In sequence from left to right—frequency versus intimacy, recency versus intimacy, and frequency versus recency respectively

Table 3 Mean and SD of parameters for training

	Frequency	Intimacy	Recency
Mean	0.254138	0.438586	0.779383
SD	0.422264	0.458369	0.314565

Besides, to call behavior, human interaction is dependent on other factors as well, like age, gender, social status, and community. To further dive deep into the behavioral patterns a case study can be made based on diverse groups of participants on a weekly or monthly basis across different communities [26]. This can help discover other features that can enhance the accuracy of trust prediction by a neural network. One of the primitive features that may be taken into consideration under this context is age and gender, which are vital factors of determining human behavior. Some special cases of call behavior can also exist where people might not have regular conversations with a user, resulting in less total duration, yet the duration of a single call may be high. This may indicate that the user has a high trust for the other person but it will be hard for a machine to realize that. To tackle this, we may consider the standard deviation of average call as an added feature to our prediction model, as outliers will have extremely high deviation values which will demarcate them from normal records.

Challenge 2: Scaling up

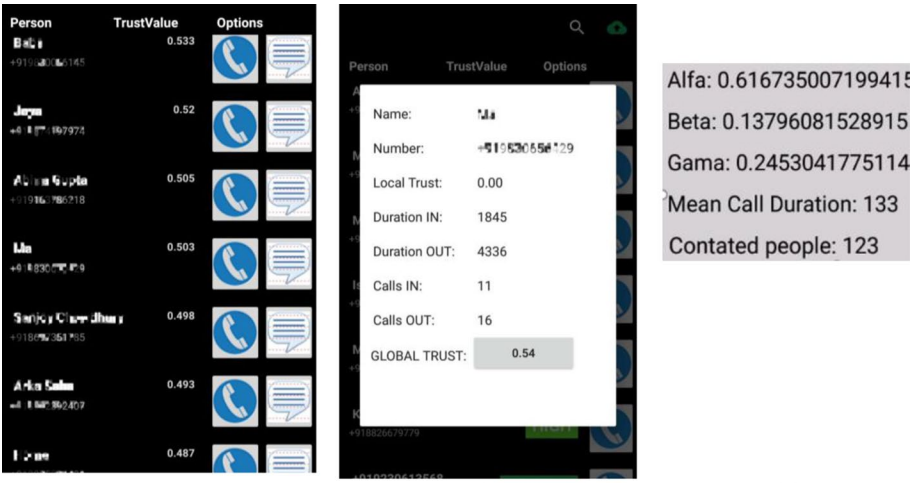


Fig. 10 **a** Decimal Trust Calculation using Method 1, **b** local and Global Trust of a random person, **c** calculated values of α, β and γ for an arbitrary user by Method 1

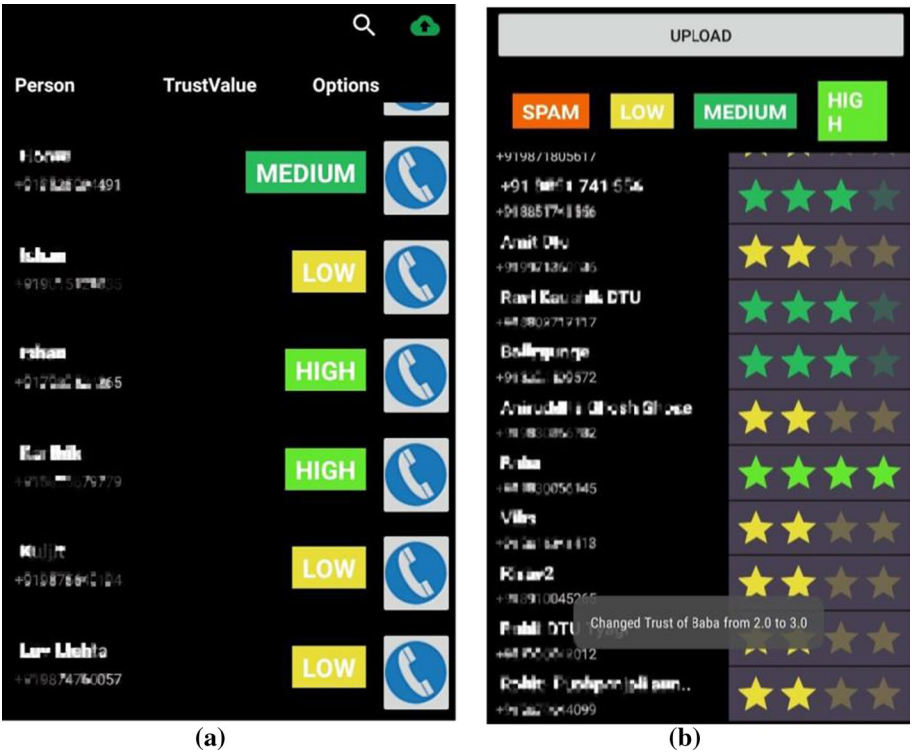


Fig. 11 **a** Categorical Trust calculation using method 2, **b** user corrected feedback on predicted Trust ratings to be uploaded on the server

To get more data, we need a larger audience and a broader spectrum. Online services like Facebook Messenger and WhatsApp can be alternative services to voice calls, which might be preferred by some people over voice calls. In these cases, the data from these services also may be incorporated with equal weightage in determining the trust of a person, along with voice call features. Various performance improvements can be made in the application by identifying memory leaks [24] and optimizing retrieval methods and server interaction to support a huge community of users, and analysis using Big Data Computing to scale the application to a production environment and thus build a global community [27] of participants for further study.

Challenge 3: Role of Trust management in various research domains

Based on call interactions and user data, recommender systems may be developed using mobile call analytics. Fuzzy Trust-based recommender system and Rough Fuzzy recommender system may be an interesting area of future trust management, with understandable privacy policies [25, 27, 28]. Hash Key-based Blockchain is considered as one of the potentially disruptive technologies in Industry 4.0. Fog, Edge, and Dew [23, 28] computing-based secure IoT applications can be managed efficiently by Mobile Trust. HetNet Trust management using Blockchain may open a new research avenue in 5G network.

7 Conclusion

The trust model is comprehensively dependent on behavioral patterns along with the basic feature attributes. However, the trust model also has to assume that all user ratings and feedback are genuine and valid. Our research has contributed to the significance of normalizing features based on these behavioral attributes and also realized that the pattern for trust determination is more complex than just a linear solution. Categorizing trust values has made the analysis more understandable and also globally applicable, without having any user-generated bias. Gathering feedback data from crowdsources to validate the label of a record, as well as to rectify incorrect labels will ensure even better predictions, with time. Having a copy of the perceptron model, in every device, which can be trained based on user feedback can enable one to have an adaptive, personal trust generator with them. Moreover, the generated data can be used for spam detection which can be a proposed extension of the current project, considering the datasets already available.

References

1. Bok, K., Yoon, S., & Yoo, J. (2018). Trust evaluation of multimedia documents based on extended provenance model in social semantic web. *Multimedia Tools and Applications,* 1–22.
2. Anakath, A. S., Rajakumar, S., & Ambika, S. (2017). Privacy preserving multi factor authentication using trust management. *Cluster Computing,* 1–7.
3. Singh, O., Singh, J., & Singh, R. Multi-level trust-based intelligence intrusion detection system to detect the malicious nodes using elliptic curve cryptography in MANET. *Cluster Computing,* 1–13.
4. Kim, M., & Park, S. O. (2013). Group affinity based social trust model for an intelligent movie recommender system. *Multimedia Tools and Applications, 64*(2), 505–516.
5. Li, H., Zhang, S., Hu, Y., Shi, J., & Zhong, Z.-m. (2017). Research of social recommendation based on social tag and trust relation. *Cluster Computing,* 1–11.

6. Santhi, G., Maria, S., Anna, W., & Sengutuvan, P. (2013). Fuzzy rule based novel approach to spam filtering. *International Journal of Computer Applications (0975 – 8887),* 71(14).
7. Chakraborty, S., & Mondal, B. (2012). Spam mail filtering technique using different decision tree classifiers through data mining approach—a comparative performance analysis. *International Journal of Computer Applications (0975–888), 47,* 16–26.
8. Ramchoun, H., Idrissi, M. A. J., Ghanou, Y., & Ettaouil, M. (2016). Multilayer perceptron: architecture optimization and training. *IJIMAI, 4*(1), 26–30.
9. Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A practical guide to support vector classification*.
10. Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). *KNN Model-based approach in classification. OTM on the move to meaningful internet systems 2003: CoopIS, DOA, and ODBASE*, pp. 986–996.
11. Chen, S., Wang, G., & Jia, W. (2013). A trust model using implicit call behavioral graph for mobile cloud computing. *Cyberspace safety and security,* pp. 387–402.
12. De, D., & De, B. C. (2018). TBFN: Trust based friend network deisgn by analyzing user's voice call and SMS pattern. *Wireless Personal Communication,* Springer.
13. Xin, L., Leyi, S., Yao, W., Zhaojun, X., & Wenjing, F. (2013). A dynamic trust conference algorithm for social network. *Eighth international conference on P2P, parallel, grid, cloud and internet computing*.
14. Oh, H.-K., Kim, J.-W., Kim, S.-W., & Lee, K. (2018). A unified framework of trust prediction based on message passing. *Cluster Computing,* 1–13.
15. Ju, C., Wang, J., & Xu, C. (2018). A novel application recommendation method combining social relationship and trust relationship for future internet of things. *Multimedia Tools and Applications*.
16. Adali, S., Escriva, R., Goldberg, M. K., Hayvanovych, M., Magdon-Ismail, M., Szymanski, B. K., Wallace, W. A., & Williams, G. (2010). Measuring behavioral trust in social networks. *IEEE international conference on intelligence and security informatics*.
17. Kim, M., & Park, S. O. (2013). Trust management on user behavioral patterns for a mobile cloud computing. *Cluster Computing, 16*(4), 725.
18. Zhao, B., Xiao, C., Zhang, Y., Zhai, P., & Wang, Z. (2018). Assessment of recommendation trust for access control in open networks. *Cluster Computing*.
19. Hosseini, M., Angelopoulos, C. M., Chai, W. K., & Kundig, S. (2018). Crowdcloud: A crowdsourced system for cloud infrastructure. *Cluster Computing,* 1–16.
20. Chaisamran, N., Okuda, T., Blanc, G., & Yamaguchi, S. (2011). Trust-based VoIP spam detection based on call duration and human relationships. *IEEE*.
21. Singh, S. P., & Yadav, A. Study of K-Means and enhanced K-Means clustering algorithm. *International Journal of Advanced Research in Computer Science*.
22. Aitkin, M., & Foxall, R. (2003). Statistical modelling of artificial neural networks using the multi-layer perceptron. *Statistics and Computing, 13,* 227–239.
23. Jia, G., Chen, R., Wang, D.-C., Tsai, J. J. P., & Al-Hamadi, H. (2019). Trust-based IoT cloud participatory sensing of air quality. *Wireless Personal Communications,* 1–14.
24. Riganelli, O., Micucci, D., & Mariani, L. (2018). From source code to test cases: A comprehensive benchmark for resource leak detection in Android apps. *Software: Practice and experience*, pp. 1–9.
25. Dev, J., Das, S., Rashidi, Y., & Camp, L. J. (2019). Personalized WhatsApp privacy: Demographic and cultural influences on Indian and Saudi users: SSRN.
26. Cheok, A. D., Edwards, B., & Muniru, I. (2017). Human behavior and social networks. In R. Alhajj & J. Rokne (Eds.), *Encyclopedia of social network analysis and mining*. New York, NY: Springer.
27. Ramjee, P., & Ruggieri, M. (2016). Special Issue on Human bond communications (HBC). *Wireless Personal Communications, 88*(1), 1–4.
28. De, D. (2016). *Mobile cloud computing: Architectures, algorithms, and applications*. CRC Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Arka Bhowmik is a Masters student at Delhi Technological University, New Delhi, India in the field of Computer Science and Technology. He has expertise in android development, web development and data science. His current research interests include Deep Learning, Machine Learning and Social Networks.

Dr. Debashis De (M'13-SM'15) is a Professor and Director of the Department of Computer Science and Engineering of Maulana Abul Kalam Azad University of Technology, India and Adjunct Research Fellow of University of Western Australia, Australia. His research area includes energy and latency optimization in mobile cloud computing. He has received Young Scientist award both in 2005 at New Delhi and in 2011 at Istanbul from International Union of Radio Science, H. Q., Belgium. His email id: dr.debashis.de@ieee.org.

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 11, November-2020

Multi-Output Fast Wireless Charging for Electric Vehicle

Electric Vehicle Charging

Imroz Khan
Student of Electrical Engineering
Delhi Technological University
Delhi, India

Anjali Kumari
Student of Electrical Engineering
Delhi Technological University
Delhi, India

Hemant Sharma
Student of Electrical Engineering
Delhi Technological University
Delhi, India

Prof. Prem Prakash
Faculty of Electrical Engineering
Delhi Technological University
Delhi, India

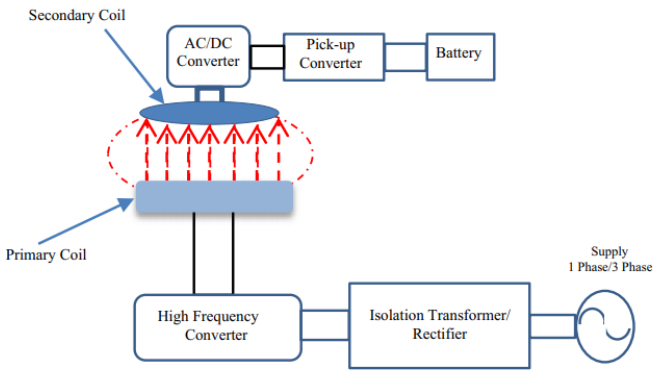
***Abstract* — To ensure a greener and cleaner ecology it is necessary to progressively reduce carbon emission from our earth. This can be typically achieved by a substantial cut in fossil fuel-run vehicles and shifting to electric cars which consume generated electricity as fuel, providing fewer emissions than a conventional vehicle, and will be zero-emission if used electricity comes from a renewable source. Electric vehicles are typically powered by batteries, which need to be recharged from generated electricity from the grid. Batteries can be charged either by plug-in methods or by wireless charging. As charging the battery pack requires quite a lot of time throw wireless charging. This research explores the Proposed multi-output charging method for the fast charging of the electric vehicle in this method we will divide the battery pack into four stacks and will charge every battery stack with a different output terminal (multi-output) terminal by using three winding transformers. This system can help in reduce the charging time of the electric vehicle. The conclusion has been made on the basis of calculation study of multi-output system.**

Keywords—Wireless charging system; Electric vehicle; Multi-output system; charging time.

INTRODUCTION OF WIRELESS CHARGING

Electric vehicles are observed as an alternative option in response to the depletion of resources. In order to increase the use of EVs in daily life, practical and reliable methods to charge batteries of EVs are absolutely important. The wireless charging has started an alternative methodology for charging electrical gears. As charging of battery throw wireless technology for Electric Vehicles (EVs) arise numerous issues like the power pad, coil arrangement plan's and most significant is charging time for batteries of an electric vehicle, power converters for high recurrence power change, and also electromagnetic field securities. All the above challenges are extremely important to research and to develop the solutions. In this holograph, our focus is on the fast charging analysis of the electric vehicles throw a multi-output charging arrangement to minimize the charging time for an Electric vehicle.

Inductive wireless Inductive wireless charging uses the principle of mutual induction by linking the transmitter and receiver to a time-varying magnetic field. The elements of the wireless charging system include a low-frequency ac / dc converter, high-frequency dc / ac converter, transmitter and receiver coils, secondary side high-frequency ac / DC The power is being transferred from the grid to the battery wirelessly through an ac/dc converter which converts grid supply to DC. This dc is then converted to high-frequency ac through the dc/ac converter. The primary and secondary coils are coupled to transfer the power at high frequency. Secondary side high-frequency ac/dc converter sends the current to variable dc which is regulated and fed to the battery for charging. The power transfer efficiency can be improved by combining the LCL or LC resonant circuit's charged using inductive power transfer can be charged in the static mode as well as dynamic mode. Both of the cases need to have proper alignment between the transmitter and receiver.



TYPICAL EV WIRELESS CHARGING SYSTEM

THEORY OF THE PROPOSED IDEA

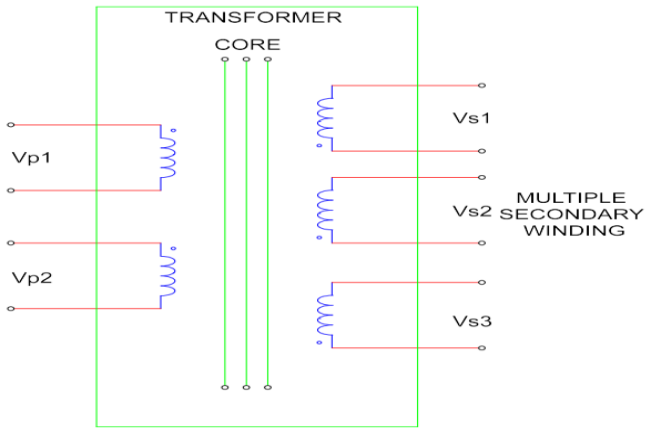
- In this proposed idea we are using a multiple winding transformer concept for fast wireless charging. Multi tuned circuit comprise of one essential winding in primary and at least two secondary winding. The

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 11, November-2020

fundamental freedom of this type of circuit is that it allows one to have either the main or secondary side with more than one winding.

- The principle of operation of multi tuned circuit is same as of an ordinary transformer all the calculations such as primary and secondary current, transformation ratio is all calculated the same way. But in the multi tuned circuit, we must pay exceptional attention to the voltage polarities of each coil wrapping, and also the dot convention reflects the winding's positive or negative polarities when we interface the two windings together.
- The multi-tuned circuit also known as the multi-coil transformer which consists of more than one winding either on the primary or secondary wound on a common laminated core. A multi-tuned circuit can be used as the step up and step-down circuit or the combination of both. This multi-tuning system may also include multiple secondary windings with a different voltage or a different current level output. Since the tuned circuit works under the mutual induction principle, each individual winding of a multiple winding transformer is firmly support's the same volt / ampere product per turn, which means NP / NS = VP / VS with any turn's ratio between the individual coil windings being relative to the primary supply.
- A multi-coil transformer or a multi-tuned circuit is routinely used to supply different components of the electronic circuit with adjustable lower voltage levels. In power supplies and Triac switching converters, the main application of multi tuned circuits is A multi-coil transformer / multi-tuned circuit may also have a number of distinct secondary windings that are electrically isolated from each other, as both primary and secondary windings are electrically isolated from each other. In that instance, every individual secondary coils will produce a voltage that refers to the number of turns of their coils.

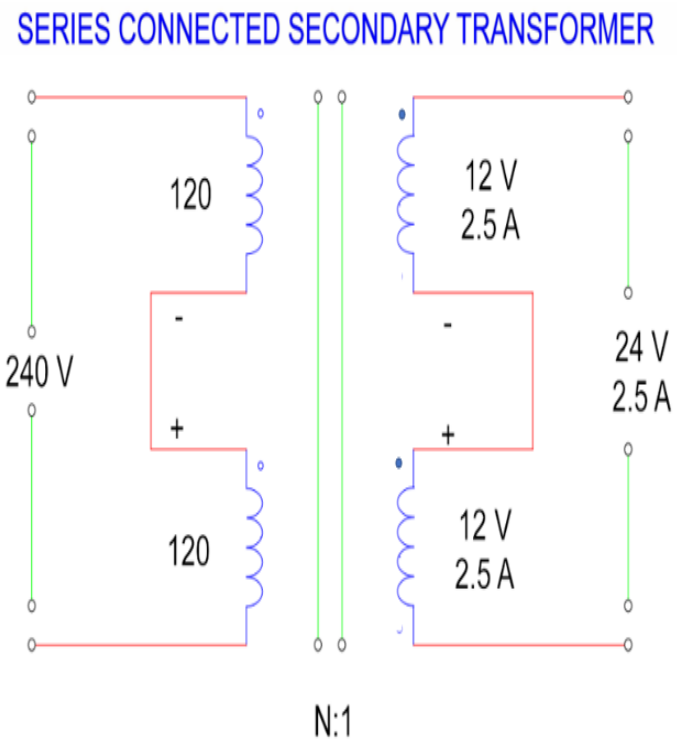


MULTI-WINDING TRANFORMER CONCEPT

Fig-1 is a typical multiple winding transformer in this case, which has multiple secondary windings with different voltage levels. The primary winding can be used separately or linked together to operate the higher supply voltage transformer

The secondary winding may be connected in different configurations to generate a higher voltage or current supply. It should be remembered that parallel winding connections are only possible if the two windings are electrically similar in order to be equal to their current and voltage ratings.

SERIES CONNECTED SECONDARY TRANSFORMER

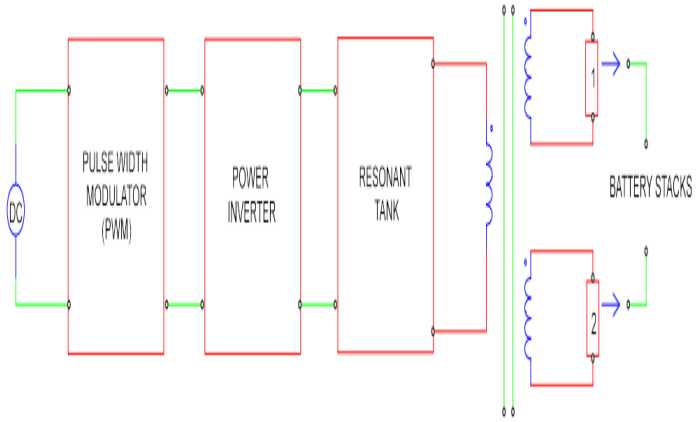


- In this example the two 120 V rated primary winding are connected together in series across a 240 V supply as the two windings are identical. Half the supply voltage normally 120 V is dropped across each winding and the same primary current flows through both.
- The two secondary winding rated at 12 V, 2.5 A each is connected in series with the secondary terminal voltage which is the sum of two individual winding voltage giving 24 V.
- As the two windings are connected in series the same amount of current flows through each winding, as the secondary current is the same in both secondary windings which are 2.5 A. so in a series-connected secondary, the output is rated at 24 V and 2.5 A

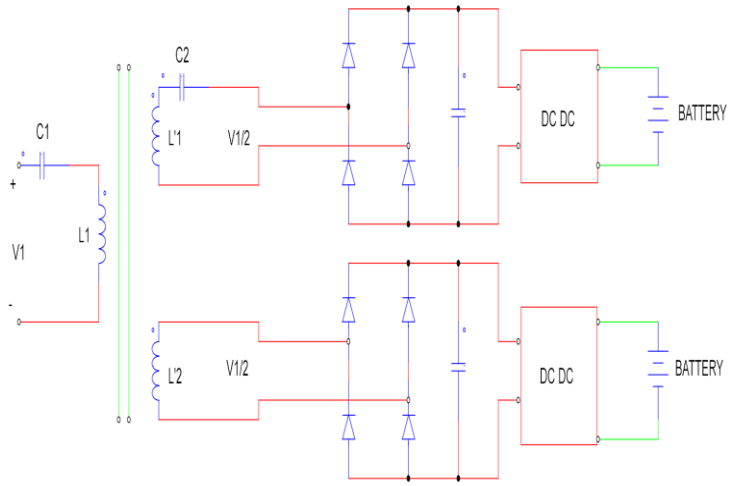
(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 11, November-2020

PROPOSED MODEL



CIRCUIT DIAGRAME OF THE PROPOSED MODEL



PULSE WIDTH MODULATOR

A method of reducing the mean power supplied by a signal by effectively cutting it in discrete sections is pulse-width (PWM) or pulse-duration (PDM). By switching between the supply and load, the average voltage value (and current) supplied to the load is regulated at rapid speed. The higher the total power supplies to the load, the longer the switch is turned on, the higher the total power supplied to the load. The PWM switching frequency must be high enough to not affect the load.

POWER INVERTER

A power converter or inverter is a power electronic unit or circuit that converts DC into AC. Each system or circuit depends on the input voltage, output voltage, frequency, and general power management. There is no power-producing inverter; the DC source provides electricity. A power inverter may be fully electrical, or a combination of mechanical and electronic effects (such as a rotary device). Static inverters do not use moving parts during the conversion process.

RESONANT TANK

The electrical circuit which mainly consists of an inductive circuit, defined by letter L, and a capacitor, represented by letter C, linked together; this type of circuit is often referred to as the resonant circuit, a tank circuit, or a tuning device. The circuit acts as an electrical resonator, a tuning fork analog, that stores oscillating energy at the resonant circuit frequency. LC circuits are either used for signal generation in a specific frequency or for signal selection from a more complex signal at a certain frequency; this purpose is known as a bandpass filter. The LC circuit is an idealized model since it assumes that there is no energy dissipation due to resistance. The aim of the LC circuit is typically to oscillate with minimal damping making the resistance as low as possible. though there is no practical circuit without losses

CALCULATION

CASE – 1 WITHOUT MULTI-OUTPUT CHARGING
ASSUMING
RATING OF BATTERY – 120 A/HR
CHARGING CURRENT – 12 A
TIME OF CHARGING - 120/12 = 10 HOURS
NOW, IF THERE ARE 2 STACKS OF BATTERY PACK IN OUR PROPOSED MODEL.
FOR THE FIRST STACK OF BATTERY PACK
A/HR RATING = 60 A/HR
CHARGING CURRENT = 12 A
TIME OF CHARGING = 60/2.5 = 5 HOURS
FOR THE SECOND STACK OF BATTERY PACK
A/HR RATING = 60 A/HR
CHARGING CURRENT = 12 A
TIME OF CHARGING = 60/2.5 = 5 HOURS
THIS METHOD WILL CHARGING OUR BATTERY IN ALMOST HALF THE TIME TAKEN IN NORMAL CASE (WITHOUT USING MULTI-OUTPUT CHARGING)

CASE- 2- WITHOUT MULTI-OUTPUT CHARGING TAKING HIGH CHARGING CURRENT
RATING OF BATTERY – 120 A/HR
CHARGING CURRENT – 20 A
TIME OF CHARGING - 120/20 = 6 HOURS
NOW, IF THERE ARE 2 STACKS OF BATTERY PACK IN OUR PROPOSED MODEL.
FOR THE FIRST STACK OF BATTERY PACK
A/HR RATING = 60 A/HR
CHARGING CURRENT = 20 A
TIME OF CHARGING = 60/20 = 3 HOURS
FOR THE SECOND STACK OF BATTERY PACK
A/HR RATING = 60 A/HR
CHARGING CURRENT = 20 A
TIME OF CHARGING = 60/20 = 3 HOURS
THIS METHOD WILL CHARGE OUR BATTERY IN ALMOST HALF THE TIME TAKEN IN NORMAL CASE (WITHOUT USING MULTI-OUTPUT CHARGING)

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 11, November-2020

The main purpose of this model is that instead of charging the whole battery pack at a time we instead divided the battery pack into number of stacks and thus charging each individual stack with Multi winding this will reduce the charging time as shown in the calculation.

DISCUSSION AND CONCLUSION

This paper contains a review of the Inductive wireless power transfer technologies and applications in transportation. In this paper, we have discussed the new concept and design of the multi-output wireless charging of the electric vehicle. The proposed system can independently charge several output batteries of EVs at the same time with the help of a multi output winding. As the future belongs to electric vehicles and the main problem which exists is the charging time of the electric vehicle thus by using this topology, we can actually reduce the charging time of Electric vehicles.

REFERENCES

[1] Van-Binh Vu, Mohamed Dahidah “Multiple output inductive charger for electric vehicle” IEEE TRANSACTIONS ON POWER ELECTRONICS, VOL. 34, NO. 8, AUGUST 2019.
[2] Kishore Naik Mude, ”Battery Charging Method for Electric Vehicles: From Wired to On-Road Wireless Charging” Chinese Journal of Electrical Engineering, Vol.4, No.4, December 2018.
[3] Miss. Shital R. Khutwad, Mrs. Shruti Gaur “Wireless Charging System for Electric Vehicle” International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016
[4] Multi-winding transformer, Multi-coil winding https://www.electronics-tutorials.ws
[5] LC circuit, Wikipedia,(28 oct 2020).
[6] Pulse-width modulation, Wkipedia (30 oct 2020.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

Nanomaterials Based Biosensors for COVID-19 Detection-A Review

Sakshi Sharma, Sonakshi Saini, Maya Khangembam, Vinod Singh*

Department of Applied Physics, Delhi Technological University, New Delhi, India-110042
* vinodsingh@dtu.ac.in

***Abstract*—This review paper discusses the properties of nanomaterials, namely graphene, molybdenum disulfide, carbon nanotubes, and quantum dots for unique sensing applications. Based on the specific analyte to be detected and the functionalization techniques that are employed, some noteworthy sensors that have been developed are discussed. Further, biocompatible sensors fabricated from these materials capable of detecting specific chemical compounds are also highlighted for COVID-19 detection purposes, which can aid in efficient and reliable sensing as well as timely diagnosis.**

***Index Terms*— graphene, quantum dots (QDs), sensors, biocompatible, COVID-19**

© IEEE 2020. This article is free to access and download, along with rights for full text and data mining, re-use and analysis
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

I. Introduction

Nanomaterials such as graphene, transition metal dichalcogenides and carbon nanotubes have attracted research efforts steadily over the last decade for numerous applications, as a result of their unique properties. For use in sensors, the superior electronic[1], optical[2], mechanical[3] and structural[4] properties of 2D substances and thin films, such as a high surface-to-volume ratio, enhanced flexibility, optical transparency, presence of multiple reaction sites, and high surface sensitivity, all contribute towards efficient sensing and detection of a range of analytes and molecules[5, 6].

Additionally, quantum dots are also beginning to be increasingly incorporated in sensors and imaging owing to their singular properties – large Stokes shift, narrow emission band, large molar absorption coefficient, long excited state lifetimes, high fluorescence quantum yields, superior resistance to chemical degradation, photobleaching and a high two-photon absorption cross section.

In this review, we have explored advances in sensors fabricated using these nanomaterials and discussed the application of these sensors for biosensing and, in particular, for detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus.

II. Materials For Biosensors

A. Graphene

Graphene is a zero-bandgap semiconductor and possesses a low defect crystal structure. Owing to its pi-bonds, graphene exhibits extremely superior electrical conductivity by transforming chemical signals into electric current through transduction. It also depicts enhanced mechanical properties, including high tensile strength, flexibility, and elastic modulus. In spite of being atomically thin, it exhibits high absorbance, and is also thermally as well as chemically stable. Moreover, graphene can be incorporated with other polymer nanocomposite materials to obtain desirable properties for detecting trace amounts of specific compounds or harmful toxic elements such as heavy metals or industrial chemicals[7, 8].

Based on the specific target molecule or gas to be detected by the sensor, graphene can be functionalized with different compositions of metal oxides. The ease of availability of materials, combined with the advantage of graphene's biocompatibility, have led to several graphene-based biosensors being fabricated, for the detection of specialized compounds[9], such as paracetomol[10], ascorbic acid[11], aromatic isomers[12], prostate specific antigen(PSA)[13], and hemoglobin(Hb)[14].

Sun et al. fabricated an electrochemical DNA biosensor for the detection of MON810, a genetically modified gene sequence via electrostatic adsorption through an electrochemical reduced graphene (ERG) film with a modified carbon ionic liquid electrode (CILE). The sensor displayed high sensitivity and the presence of the ERG film contributed towards increasing the adsorption of the probe ssDNA, which further enhances the response of methylene blue (MB), an electrochemical indicator, which can interact with DNA[15].

To harness the electrocatalytic energy of functionalized graphene sheets (FGS) for sensing glucose, Wu et al. reported a sensor consisting of a bionanocomposite film prepared by electrodeposition of modified platinum nanoparticles on the FGS (which was initially dispersed in a solution of chitosan and glassy-carbon electrode) and subsequent modification of the resultant film with enzymes. Lastly, glucose oxidase was immobilized on the film. The resultant biosensor depicted good sensitivity and responsivity, high reproducibility and long-term stability, thereby showing great promise for rapid and accurate glucose monitoring. Moreover, this fabrication method can also be extended for other types of graphene-based biosensors[16].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

TABLE I
GRAPHENE BASED BIOSENSORS

S. NO.	ANALYTE	BIOSENSING ELEMENT	ADVANTAGES AND SENSOR CHARACTERISTICS	REFERENCE
1	Paracetamol	Glassy carbon electrode modified with carbon	High sensitivity and recovery, detection limit of 3.2 x 10^{-8} M and 5.2% reproducibility	10
2	Ascorbic acid	Graphene doped carbon paste electrode	Low overvoltage, good current response and sensitivity, high response rate of 5s, and detection limit of 7 x 10^{-8} M	11
3	Hydroquinone (HQ) and Catechol (CC)	Reduced graphene oxide nanosheet on glassy carbon electrode	Enhanced sensitivity, high reproducibility and stability, detection limits of 0.2, 0.1 μM for determination of HQ and CC respectively	12
4	Prostate Specific Antigen (PSA)	Graphene sheet - cobalt hexacyanoferrate nanoparticle - 1-pyrenebutanoic acid, succinimidyl (GS–CoNP–PBSE) on glassy carbon electrode	Label-free detection, high electroactivity, good stability, reproducibility (6.7% standard deviation) and selectivity (<8% current variation due to interfering substances)	13
5	Hemoglobin (Hb)	Graphene nanosheets (PDDA-G), poly (diallyldimethylammonium chloride), and room temperature ionic liquid (RTIL) nanocomposite	Superior electrocatalytic activity, biocompatibility and solubility, linear range from 0.2 - 32.6 μM, detection limit of 0.04 μM at 3σ	14
6	ssDNA (for detection of transgenic maize MON810)	Modified carbon ionic liquid electrode (using electrochemically reduced graphene)	Good stability, sensitivity and conductivity, simple fabrication procedure, low detection limit of 4.52×10^{-12} mol/L (3σ)	15
7	Glucose	Glucose oxidase/Pt/functional graphene sheets/chitosan (GOD/Pt/FGS/chitosan) bionanocomposite film	Good sensitivity, reproducibility, stability and response, low detection limit of 0.6 μM	16

These graphene-based sensors have been summarized in table I for easy reference.

B. Transition Metal Dichalcogenides (TMDs)

Among two dimensional materials, transition metal dichalcogenides (TMDs) garner significant research interest as some of these materials possess the desirable quality of having tunable bandgaps (indirect to direct), which resulting in some unique properties such as photoluminescence. In the family of TMDs, molybdenum disulfide (MoS_2), molybdenum diselenide ($MoSe_2$) tungsten disulfide (WS_2) and tungsten diselenide (Wse_2), have been extensively studied for use in sensors, with special focus on MoS_2 due to its superior conductivity, fast electron transfer rate, and easy availability[17]. These advantages allow MoS_2 to be efficiently used for many diverse applications, from energy storage and flexible electronics, to biological applications owing to its low toxicity and stable nature[18].

The structural defects present in MoS_2 can be used for surface modifications or functionalization, and mono or few layer MoS_2 nanosheets depict confinement effects, all of which result in enhanced properties for use in sensors. MoS_2-based biosensors have been fabricated and reported to detect biomolecules such as DNA[19] and lactate[20]. Based on the type of MoS_2, it may be semi-conducting with a tunable direct bandgap (2H) or metallic (1T)[21].

Lee et al. reported an MoS_2 based label-free biosensor utilizing a simplified dielectric-free design for the detection of prostrate specific antigen (PSA)[22]. Huang et al. presented an electrochemical biosensor fabricated from MoS_2 nanocomposites[23].

Thus, the unique properties of MoS_2 combined with efficient fabrication methods offer exciting opportunities for application in optical, electronic, and electrochemical biosensors.

These MoS_2-based biosensors have been summarized in table II for easy reference.

C. Carbon Nanotubes (CNTs)

CNTs, another nanomaterial made of carbon, have been extensively studied for applications corresponding to a broad

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

TABLE II
MoS_2 BASED BIOSENSORS

S. NO.	ANALYTE	BIOSENSING ELEMENT	ADVANTAGES AND SENSOR CHARACTERISTICS	REFERENCE
1	DNA	Nano-MoS_2 modified carbon paste electrode(nanoMoS_2/CPE)	Label-free detection, high electrochemical activity, detection limit of 1.9 x 10^{-17} M, good sensitivity, reproducibility and stability	19
2	Lactate	GC/MoS_2/Lox (MoS_2 nanosheets, exfoliated onto a glassy carbon electrode, along with lactate oxidase enzyme)	Linear range from 0.056 - 0.77 mM, sensitivity of 6.2 μAmM^{-1}, detection limit of 17 μM and reproducibility (RSD) = 4.7%	20
3	Prostate Specific Antigen (PSA)	Multilayer MoS_2 FET	Oxide-free simple design, label-free detection, enhanced sensitivity,	22
4	BPA (bisphenol A)	MoS_2 and chitosan-gold nanoparticles modified electrode	Enhanced sensitivity, long-term stability, selectivity and repeatability (RSD=2.53%), detection limit of 5.0 × 10−9 M	23

range of fields due to their superior properties. The C-C bond makes them the stiffest and most durable fiber, whether of a one atom thickness (SWCNTs) or consisting of several layers of graphite (MWCNTs). CNTs are also thermally stable and may be metallic or semiconducting based on the fabrication parameters[24]. Due to the surface structure of nanotubes, the adsorbed molecules directly influence its electronic properties, and this sensitivity is an essential characteristic for a sensor[25].

CNT-based biosensors usually consist of a sensitive element, which involves functionalization of the CNT with a specific biomolecule (such as proteins[26]), and a tranducer[27]. Villamizar et al. reported a bioFET consisting of a network of SWCNTs to enable rapid and sensitive detection of *Salmonella* Infantis and using appropriate antibodies, such a sensor can be extended for the application of sensing other viruses or bacteria[28].

Biosensors utilizing a combination of nanomaterials such as MoS_2/MWCNTs have also been reported to combine the individual benefits of the materials[29].

CNT sensors may utilize inter-CNT (when sensing response is due to effects originating within the tubes), intra-CNT (when effects arise at contact points across tubes) or Schottky barrier (relating to effects between the tubes and electrodes) sensing mechanism. The dominant mechanism is dependent on the strength and type of the analyte, device architecture, as well as the defects in the CNT[30].

CNT-based sensors overcome the drawback of carrier migration around perturbed regions in graphene-based sensors, due to their metallic states. Further, CNTs can be functionalized in a noncovalent manner with small aromatic molecules, polymers, or metal nanoparticles to enhance specific sensor selectivity. Alternatively, by utilizing covalent functionalization approaches, the long-term stability, robustness and reproducibility of the sensors can be enhanced[30].

These CNT-based biosensors have been summarized in table III for easy reference

D. Quantum Dots (QDs)

QDs refer to semiconductor nanostructures of sizes 1 to 10 nm confined in all three dimensions by potential barriers. Quantum confinement effect and a high surface-to-volume ratio cause QDs to exhibit optical and electronic properties that differ remarkably from their bulk counterparts. They are characterized by a broad absorption spectrum which increases in magnitude towards the shorter wavelengths, large Stokes shift and an approximate Gaussian shaped narrow emission band. Due to quantization effects, the bandgap of QDs increases with decreasing diameter. Thus, emission peak wavelength changes with QD size and material, as shown in Fig 1A. This property has been utilized in multicolor imaging and multiplex analyses.

QDs have a large molar absorption coefficient, high fluorescence quantum yields with long excited state lifetimes and hence make excellent probes. Additionally, they exhibit superior resistance to chemical degradation, photobleaching and a high two-photon absorption cross section. Their optical properties are dependent on surface chemistry -the number of dangling bonds determines quantum yields and stability. Accordingly, QDS can be functionalized with organic capping ligands as per required application.

- Traditional QDs

Traditional QDs (TQDs) consist of II-VI, IV-VI and III-V semiconductors in a core-shell structure, as shown in Fig 2B, and have been popularly researched as optical transducers based on PL activation/quenching caused by physical and chemical interactions on the QD surface.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

TABLE III
CNT BASED BIOSENSORS

S. NO.	ANALYTE	BIOSENSING ELEMENT	ADVANTAGES AND SENSOR CHARACTERISTICS	REFERENCE
1	Glucose	Bi-enzymatic CNT- poly(amidoamine) (PAMAM) dendrimer nano-composite	Mediator-free detection, high sensitivity, current response of 2200 nA mM^{-1}, rapid response (1 s), wide linear response range, from 4.0 µM - 1.2 mM, detection limit of 2.5 µM	26
2	*Salmonella* Infantis	Functionalized CNTFETs (network of SWCNTs acting as the conductor channel	High sensitivity, label-free detection, at least 100cfu/mL detected in 1hr, rapid response, similar design can be used to sense other pathogenic bacteria, or viruses if an appropriate molecular receptor is employed	28
3	Sub-femtomolar DNA	DNA probe on a gold nanoparticle and MoS2/MWCNT electrode, along with glucose oxidase	Good conductivity and signal amplification, high sensitivity and selectivity, linear range from 10 - 10^7 fM, simple fabrication and detection process	29

Förster Resonance Energy Transfer (FRET) involves the energy transfer from a donor excited with radiation to an acceptor nearby in ground state, resulting in fluorescence of the donor decreasing and the acceptor increasing. Sensors with TQDs as acceptors in FRET process have been made for monitoring of DNA using oligonucleotide-modified QDs[31], sensing thrombin[32], nucleases via gold nanoparticle conjugated QDs[33], casein kinase[34] and acetylcholinesterase[35] using CdSe/ZnS QDs, multiplex sensing using streptavidin-coated QDs and gold nanoparticles[36] and identifying recombinant proteins with nickel-NTA (Nitrilotriacetic acid) conjugated CdT/CdS QDs[37]. Common nucleic acid detection methods involve FRET processes in QD-fluorophore systems, QD-QD systems, QD fluorescence quenching or target and QD-conjugate probe hybridization[38].

Similarly, TQDs as acceptors in BRET (Bioluminescence Resonance Energy Transfer) systems have been exploited for biosensing, such as glucose detection via CdTe QDs[39] and nucleic acid using carboxylated QDs[40], as well as in in vivo bioimaging applications[41]. Wang et al. developed a $CuFe_2O_4$ nanoparticles and ZnO composite for detection of toxic methylene blue molecules[42].

Apart from their use in optical transducers, TQDs find use in photoelectrochemical and gas sensors. Stretchable and humidity-resistant NO_2 gas sensor using PbS colloidal QDs (CQDs) and graphene electrodes[43] and using SnS_2 QD/graphene nanostructures[44] have been reported. H_2S gas sensors fabricated using SnO_2 CQDs - multiwalled carbon nanotubes (MWCNTs) composite[45] and NO_2 and ethanol gas sensors using monoclinic WO_3 QDs[46] have been reported. Photoelectrochemical sensors using TQDs are based on change in photocurrent due to chemical changes by enzymes, QD-protein interactions and binding reactions between biomolecules[47]. Reported applications include photoelectrochemical detection of acetylcholine[48], oxygen[49], epinephrine[50], and hydrogen[51, 52] to name a few. Although the use of TQDs in a range of biosensing and bioimaging applications has been explored, their cytotoxicity has been a cause of concern. Observed cytotoxic effects have been attributed to release of Cd^{2+} ions, aggregation of nanoparticles and presence of surface ligands that are cytotoxic. To overcome the issue of toxicity alternative QDs using InGaP, InP, carbon dots, graphene quantum dots etc. are being explored[53].

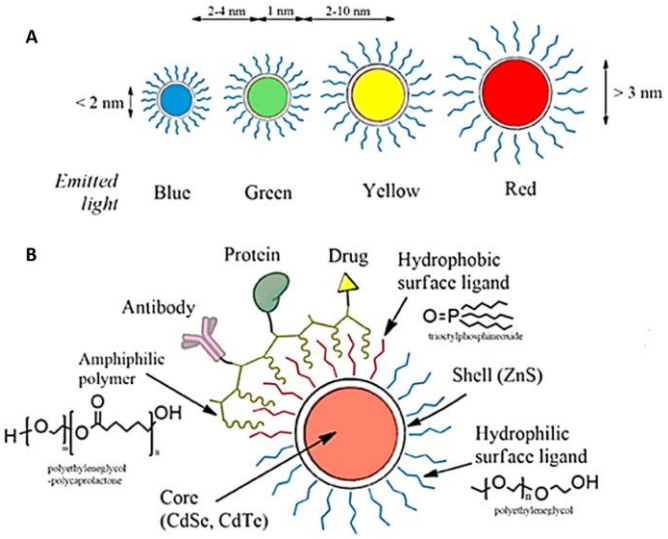


Fig.1. A) With decrease in size of QDs the emission band peak shifts towards shorted wavelengths. B) Quantum dots consist of a core-shell structure. The surface maybe coated with a different kind of ligands for linking with molecules. Reprinted from Maysinger, D., Ji, J., Hutter, E., & Cooper, E. (2015). Nanoparticle-Based and Bioengineered Probes and Sensors to Detect Physiological and Pathological Biomarkers in Neural Cells. Frontiers in Neuroscience, 9. doi:10.3389/fnins.2015.00480, an open access source.

- Carbon Dots (CDs)

CDs are QDs with a carbon core and a functional group shell. They exhibit a number of unique properties, inheriting the excellent optical and electronic properties of traditional QDs with desirable inherent properties such as high biocompatibility and low toxicity.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

Fluorescent CD sensors use direct analyte-CD interactions, post functionalization of CDs and integration with other molecules. CD sensors for detection of iron, copper and mercury ions[54], catecholamines, dopamine and ascorbic acid[55], proteins and DNA via PL quenching have been developed. Detection of folate receptor positive cancer cells using folic acid - CDs probe[56], *Escherichia coli* has been detected using mannose-modified fluorescent CDs through the binding of the lectin units of the bacteria and mannose on the CDs and thus tagging the bacteria[57]. Similarly, papaya sourced water-soluble CDs for detection of the bacteria via the same mechanism has been reported[58].

Work has also been done on electrochemical and electrochemiluminescence sensor. Louleb et al. used N-CDs for dopamine detection in human fluids (serum and urine)[59]. Hydrogen bonding between the dopamine -NH^{3+} moiety and the surface ligands of the N-CDs caused quenching in the fluorescence of the N-CDs, the extent of which was measured and used to determine
the concentration of dopamine. CD-nanomaterial combination for epinephrine, glucose, insulin, acetylcholine detection among other biomolecules has also been reported[60].

- 2D-QDs

Two-dimensional quantum dots consist of 2D-nanomaterials which consist of graphene, transition metal dichalcogenide (TMD), transition metal oxides, etc. reduced in lateral dimensions and exhibit novel properties along with most properties of their larger 2D form.

TABLE IV
QD BASED BIOSENSORS

S. NO.	ANALYTE	BIOSENSING ELEMENT	MECHANISM	REFERENCE
1	DNA	Thiolated oligonucleotide modified mercaptopropionic acid stabilized CdSe/ZnS QDs	Fluorescent quenching of QDs due FRET process from QDs to dye molecules integrated by telomerase in DNA undergoing replication/telomerisation	31
2	Proteases	Gold nanoparticles tethered QDs via peptide sequence corresponding to target protease	Quenched fluorescent AuNP-QDs show strong emission due to disabled FRET process on account of breaking of peptide chain by target protease	33
3	Recombinant proteins	Carboxylated QDs covalently bound to primary amine group attached Ni^{2+} and NTA ligand	Detection by binding of fluorescent Ni-NTA QDs with histidine in the his-tagged proteins	37
4	Glucose	*Luciferease* enzyme-glucose binding protein (GBP) conjugated CdTe QDs	GBP separating QD and enzyme binds with glucose and shortens donor acceptor distance resulting in increase in BRET emission	39
5	Dopamine	Carbon dots	Quenching of fluorescent CDs due electron-hole nonradiative recombination	55
		Nitrogen doped CDs	Quenching due to H-bonding between the dopamine-NH^{3+} moiety and the surface ligands of the N-CDs	59
6	Cancer cells	Folic acid - CDs probe	Recovery of fluorescence due to folic acid leaving CD surface to combine with folate receptor on cancer cell	56
7	*Escherichia coli*	Mannose-modified fluorescent CDs	Tagging of bacteria by binding of the lectin units of the bacteria and mannose on the fluorescent CDs	57
8	microRNA	GQDs as horseradish peroxidase immobilization platform	Increase in current via HRP-GQD catalyzed electrochemical reduction of 3,3′,5,5′-tetramethylbenzidine (TMB)	61
9	Metronidazole (MNZ)	Molecularly imprinted polymers (MIP) on GQDs and graphene nanoplatelet modified glassy carbon electrode (GCE)	MNZ accumulation on GQD-MIP recognition sites on electrode surface	62
10	Peroxide	Graphene QD assembled on gold electrode	Measurement of current due to reduction of H_2O_2 as it undergoes electrocatalysis	64

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

Functionalization of 2D-QDs allow for enhancing specific properties and thus for application in various sensors. Chemiluminescence sensors for uric acid and 4-nitrophenol detection, electrochemical sensors for microRNA and caffeic acid detection, and electrochemiluminescence sensors for detecting adenosine triphosphate, metal ions, microRNA and antigens based on 2D-QDs have been developed[61].

Electrochemical sensors based on graphene QDs explored include detection of metronidazole via molecularly imprinted polymers on GQDs and graphene nanoplatelets[62] and more recently, dopamine release detection with composite of GQDs and MWCNTs[63]. An electrode based on graphene QD assembled on gold electrode was developed by Zhang et al. which was able to detect H_2O_2 by measuring the current during the reduction of H_2O_2 through electrocatalysis[64].

These QD-based biosensors have been summarized in table IV.

III. Biosensors For COVID-19 Diagnosis

The first human cases of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) or COVID-19 were recorded in Wuhan, China in December, 2019 before the World Health Organization declared the global outbreak of the highly contagious virus a pandemic on March 11, 2020. [65]

Thanks to the continual medicine research traversing the spectrum, from early detection to the treatment- some even highlighting the very possibility of a disease outbreak in the years leading up to 2020 [66], we can rely on several techniques to detect COVID-19[67, 68]. These include Nucleic Acid Tests-Gene based identification of viral genomic RNA using Reverse Transcription Polymerase Chain Reaction or RT-PCR tests as the gold standard, Serological Immunoassays adopted in conjunction with RT-PCR to get very high detection rates (98.6%) [69, 70], CT Imaging[71] and Biosensors. These are delineated in Fig 2. The current testing standards, however, present certain challenges. The description of the basic working along with advantages and disadvantages of the current diagnostics tests is as follows.

RT-PCR Tests detect viral RNA. Enzymes convert RNA to DNA, which is replicated billions of times due to temperature cycles in the PCR machine, followed by binding of fluorescent markers which give a positive result if the fluorescence crosses a threshold value.

As they only detect the presence of an active virus at the time of testing, RT-PCR tests do not indicate whether a patient had a past infection and has since recovered. They require sample pre-processing for RNA extraction, removing fats, proteins and other molecules present in nasopharyngeal and nasal swab samples. Additionally, given that the virus is distributed non-uniformly along the respiratory tract, negative tests results could also mean the absence of virus at the site of

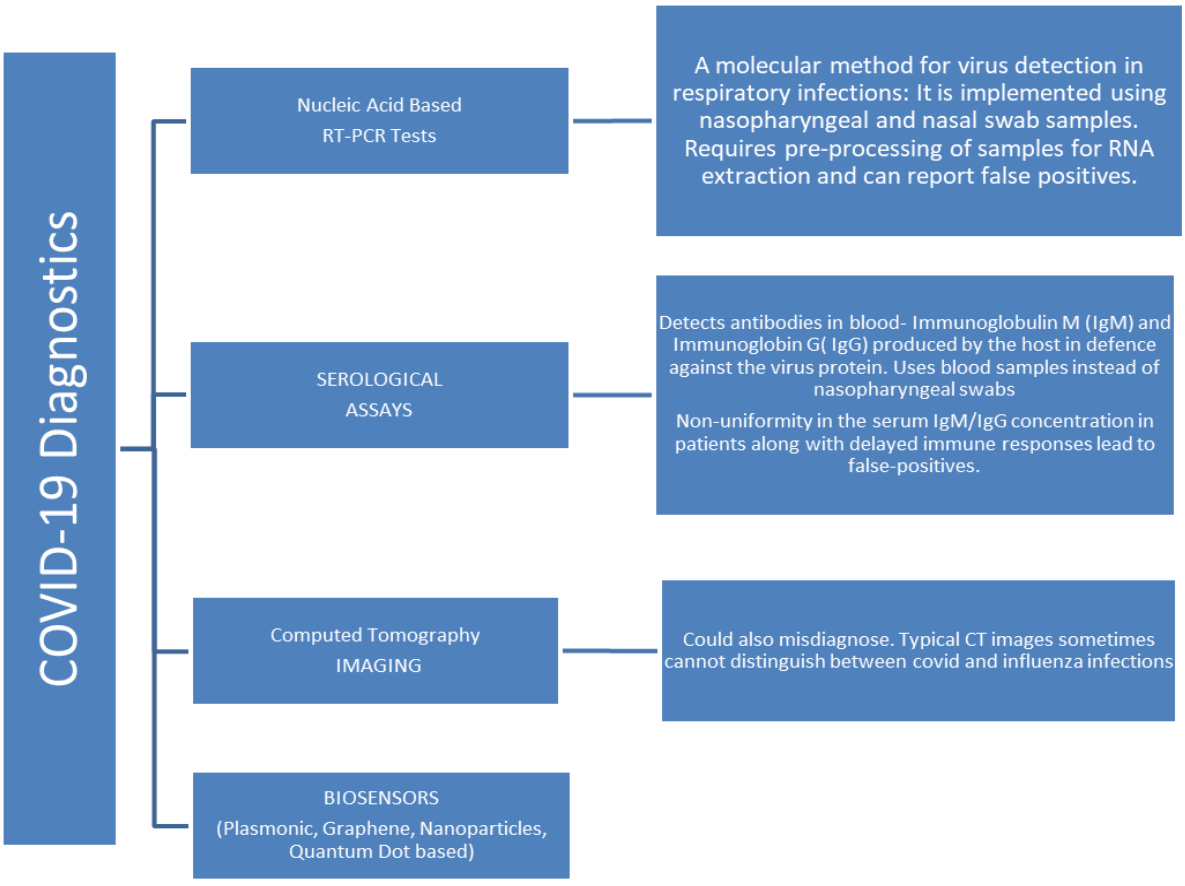


Fig.2. Methods employed for COVID-19 Diagnosis

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

sample collection, not the absence of virus altogether. [72] Samples must be sent to a specialised laboratory from the point of collection/testing site, resulting in longer response time, typically 48 hours. False negatives results arise from incorrect sampling site of swab collection or insufficient viral material in the sample.

Despite the fact that these are generally considered highly sensitive for SARS-CoV-2 detection, sensitivity has been reported as low as 59% in [73,74]. Reason for low sensitivity could be easy degradation of RNA which can be prevented by immediate frozen storage conditions. Therefore, mishandling samples could result in poor detection. During SARS-CoV-2 viral attack, the host RNA gets damaged and is released as fragments into the bloodstream, making their detection by RT-PCR challenging. Use of nanomaterials such as gold nanoparticles or fluorescent biomarkers can help isolate these RNA fragments and can overcome the detection problem by acting as signal enrichment tools. They typically have a shelf life upto 12 months. [75,76]

Lateral flow assays (LFA), or the 'Antibody' tests as they are popularly called, work by detecting a person's immune response to infection by the virus. Blood sample is mixed with a buffer liquid in the sample pad of the test strip to control the mobility. Moving across to the release pad, the sample analyte binds with antibodies conjugated with fluorescent nanoparticles, such as colloidal gold. Now containing the capturing antibody bound to the target analyte, the sample interacts with nitrocellulose membrane having multiple test lines, resulting in colorization of the appropriate test line, depending on the presence of Immunoglobulin M (IgM) or Immunoglobin G (IgG).

They present the advantage of a longer shelf life, upto 24 months [77], no need for refrigeration, easily understandable visual results, easier use for point of care use and elimination of the need for additional equipment for processing. Process from finger-stick to results takes within 30 minutes. Notably, they can help diagnose post infection conditions such as multisystem inflammatory syndrome[78].

One main research area is stemming from a shortcoming of LFA in conversion of visual data to quantitative results by digitization. Additionally, as analysis time is dependent on sample viscosity, there is no way to enhance test response using enzymes. Limit of sensitivity is introduced due to restriction on total volume of sample that is used [79]. Since research on COVID-19 virus is still in the development phase, there is no definite time period for which covid-19 antibodies can be believed to last. Thus, a negative test result does not rule out previous infection. Mild infection may not even develop detectable antibodies.

CT Scans: Chest CT Images of patients be atypical with the viral pneumonia infection images as the baseline. Images change at different stages of the disease, as recorded in many studies focusing on the manifestation of COVID-19 in lungs in the form of lobular and patchy ground glass opacities (GGOs), lesions, nodules or cavitation. [80–82]. This makes CT Scan a tool to monitor disease evolution.

CT Scans have the benefit of higher sensitivity than RT-PCR- 88% vs 59% and 98% vs 71% have been reported in [83,84]. The process is simple and equipment easily available in hospitals, making this method a potential rapid screening tool, at least for initial screening, especially in places facing shortage of RT-PCR test kits. However, images obtained from CT Scans for a COVID-19 patient may resemble those obtained in case of influenza infections, making this method fatally unreliable in some cases. Exposure to radiation in CT Scans makes it potentially inapplicable for children and pregnant women. Although it shows high sensitivity, The chest CT appearance is thought to be non-specific- the accuracy in distinguishing covid-19 infection from non covid viral pneumonia, by comparing CT features of the two, has been reported as low as 60%[85].

One challenging area for radiologists is to differentiate between COVID-19 infection and symptoms arising from other pneumonia like infections or lung disorders.

Work done by various research groups working towards nanomaterials based biosensors has reportedly addressed some, although not all, of the above mentioned drawbacks. Pros of using NM based biosensors include rapid response, enhanced selectivity and sensitivity, portability and cost-effectiveness [86, 87]. Moreover, most nanomaterials can be appropriately functionalized to obtain enhanced properties and advantages, including biocompatibility, along with producing satisfactory reproducibility of the biosensors[88].

Despite these advantages, full potential is not yet harnessed for on-site applications, equipment is not widespread for commercial use and existing logistical support exists for standard tests only. Some key attributes that serve as points of comparison while considering nanobiosensors as detection tools complementary with the standard techniques are listed in the comparative table V presented below.

In the direction of using biosensors for COVID-19 detection, there have been noteworthy developments, not least due to the fact that biosensors offer advantages of portability, sensitivity, and miniaturization as in the case of chip-based sensors along with small sample requirement. The sensing mechanism of a biosensor in general involves bio recognition elements, transducers and processor for data analysis, as shown in Fig 3. The biologically sensitive material acts like a template for detection while the transducer is responsible for converting the interaction between the bioanalyte an dits receptor into an electrical signal. Finally, the signal is appropriately filtered and amplified to be obtained as output. The Table-VI summarizes some recent advancement in sensors which have shown promise in responding to the novel coronavirus and is the subject of discussion for the following paragraphs.

Notably, the S1 spike protein is a major facilitator to the entry of the virus, making anti-S antibodies diagnostic markers[89]. S1 detection reliably indicates the presence of the virus itself [90, 91], even in the incubation period and in asymptomatic cases. A novel bioelectrochemical sensor[92]employs membrane engineered cells electroinserted with human S1 antibody as the biorecognition element for detection of S1 spike protein of SARS-CoV-2. It responds to the binding of the antigens (target analytes-S1 protein) with the inserted antibodies on the cells, resulting in changes in cellular bioelectric properties. This low cost, ultra rapid biosensor allows for differentiation between different protein concentrations. With variation in

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

TABLE V
COMPARISON OF PARAMETERS FOR STANDARD COVID-19 TESTS AND NANOBIOSENSORS

PARAMETER	RT-PCR	SEROLOGICAL IMMUNOASSAYS	COMPUTER TOMOGRAPHY SCANS	NANOBIOSENSORS
Sensitivity	93% positive in bronchoalveolar lavage fluid [70]	87.3% to 97.2% [70]	Higher sensitivity than RT-PCR- 88% vs 59% and 98% vs 71% [83, 84]	Very high sensitivity with detection limit as low as 10^{-18} M [95]
Response Time	48 hours	30 minutes	30-60 minutes	Real time detection in less than 30 seconds [123]
Reliability (Accuracy of Results)	False negatives results arising from incorrect site of swab collection or insufficient viral material in the sample	False negatives results arising from undetectable antibodies	False positive results due to non specificity: accuracy reported as low as 60% [85]	Shows promise to overcome to problem of false positives and negatives reported by the standard RT-PCR tests[107]
Portability	Test is done in specialised laboratories away from sample collection site	Lateral flow devices allow point of care testing	In specialised laboratories	Used as hand held devices for commercial POC applications

Binding of the antibodies corresponding to different domains of the S1 subunit, the biosensor could also be used detect out the associated coronaviruses [93, 94].

A Fibre-Optic Absorbance Biosensor invented by researchers from IITM and named P-FAB, employs a multimode U bent fibre optic probe to detect nucleocapsid (N) protein of SARS-CoV-2[95]from the non-invasive saliva samples[96]. The biosensor achieved a detection limit of 10^{-18} M, along with displaying benefits such as flexibility as well as sensitive and specific detection. The N protein, like the spike protein, is identified as a major structural protein of the SARS-Cov-2 with a high immunogenic activity, aiding viral RNA replication and present during early infection stages[97]. N protein from the saliva specimen with the detector antibodies (anti-N protein) in the bio-functionalised sensing region prompts a response measurable as absorbance change. A point of care device based on this biosensor has the potential for highly sensitive COVID detection [98, 99].

A novel optical biosensor designed to identify the RNA sequences of the SARS-CoV-2 [100], and aimed at real time monitoring of virus concentration in air, has been shown to be responsive towards SARS-CoV and exhibits the ability to distinguish between RNA of SARS-CoV and SARS-CoV-2, closely related viruses with slightly varying RNA sequences. It consists of gold nanoislands upon which DNA receptors complementary to the RNA sequences of the virus, are grafted. Molecules binding to the functionalised nanostructure help determine the presence of said RNA.

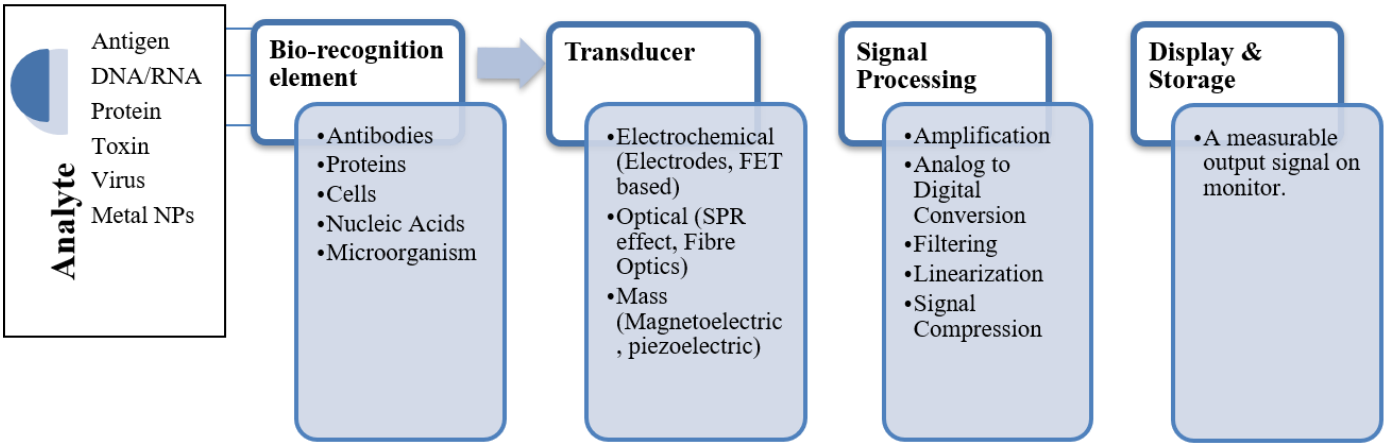


Fig. 3. Main Components of a Biosensor.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

Similar plasmonic biosensors using the combined Plasmonic Photothermal (PPT) and localized surface plasmon resonance (LSPR) effects are being explored for use in clinical diagnostics [101-105]. They are highly flexible in the choice of substrate material-from metallic nanoparticles such as gold to 2D films consisting of graphene and silicon nanowires, a number of structures easily fabricated in laboratories provide the required biocompatibility[106]. The dual function LSPR biosensor developed by Qiu et al. is particularly noteworthy due to its high selectivity and sensitivity for detecting SARS-CoV-2, with a low detection limit (0.22 pm concentration). Such a biosensor can aid in overcoming the problem of false positives or negatives reported by the standard RT-PCR tests[107].

Nanoparticles (NPs), as a consequence of their remarkable optical properties, can act as fluorescent probes for biomolecular imaging and detection [108-111]. Nanosensing systems based on Gold, Carbon, Silica and magnetic NPs in addition to Quantum Dots, have been extensively explored in relation with cellular detection and monitoring of diseases [112-120].

Functionalization using surface engineering on nanoparticles could be a useful tool to target the novel coronavirus. An exciting approach is the use of magnetic NP (MNP) based biosensors operating on the principle of giant magnetoresistance (GMR) for SARS-CoV-2 detection [120]; spin interactions between MNP and non-magnetic viral proteins are responsible for changes in resistance and therefore, magnetization. Magnetic signals detected this way contain low background noise levels due to largely non-magnetic nature of samples such as blood, serum, etc. These GMR based sensors can be translated into portable devices.

Graphene and its derivatives- Graphene oxide (GO), nano-size GO called Graphene QD, few-layer graphene, fluorographene, to name a few- are useful in their roles as components of biosensing systems. For all its interesting characteristics, Graphene is inherently inert with zero band gap, necessitating chemical functionalization using organic and inorganic molecules. Controlling its electronic behaviour by chemical modification would mean tuning the band gap, allowing its application in nanoelectronic devices. Recently, a field effect transistor (FET) based biosensing device making use of surface modified graphene sheets was introduced as a SARS-CoV-2 detector. Viral spike protein being the diagnostic antigen, the device was functionalised with receptors composed of SARS-CoV-2 spike antibody; the results demonstrated reliable detection of the antigen in cultured virus as well as clinical samples [121].

One major side effect reported for COVID-19 virus is the induction of mitochondrial reactive oxygen species (ROS) in lung cells[122]. Thus, a new approach based on sensing the level of ROS-responsible for activation of the NLRP3 inflammasome, a component of the innate immune system-has been introduced by researchers in Iran[123, 124]. It utilizes an electrochemical sensor fabricated using MWCNTs modified electrodes to measure the ROS in fresh sputum samples from patients. The notable advantages of this sensor include its ability to diagnose the viral infection in asymptomatic cases, by simply responding to the changes in ROS levels in the sputum, and the possibility of real time detection of the disease in less than 30 seconds. The tests already conducted on patients show promising results with regard to accuracy and sensitivity, providing a reasonable basis for comparison with the existing clinical tests.

TABLE VI
SOME RECENTLY DEVELOPED BIOSENSORS SHOWING POTENTIAL FOR SARS-COV-2

S. NO.	ANALYTE	BIOSENSING ELEMENT	ADVANTAGES	REFERENCE
1	S1 spike protein of SARS-CoV-2.	Membrane engineered cells electro inserted with human S1 antibody	Ultra rapid, differentiation between different protein concentrations, Detection in the incubation period/ in asymptomatic cases	74
2	Nucleocapsid (N) protein of SARS-CoV-2	Anti-N protein tagged gold NPs on U bent fiber optic probe	Minimal sample pre-processing, rapid, early detection even for low analyte concentration	77
3	RNA	Gold nano-islands grafted with which DNA receptors complementary to the viral RNA sequences	Real time monitoring of virus concentration in air	82
4	Viral proteins from blood/serum	Magnetic NP	Low background noise levels, portability	102
5	Viral spike protein	SARS-CoV-2 spike antibody immobilized graphene sheets	Highly specific and sensitive, detection of virus from clinical samples without preprocessing	103
6	Mitochondrial reactive oxygen species (ROS) in lung cells (from sputum samples)	Working electrode modified with MWCNTs	Monitoring viral infection in asymptomatic cases, real time and sensitive detection in less than 30 seconds.	105

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

IV. Challenges Ahead

A biosensor requires many capabilities such as high specificity and high sensitivity so that it is able to identify the right analyte in the midst of any other similar contaminants or analytes present. It must be non-toxic to allow use in biological systems and disposable to avoid transfer of infectious diseases. Other requirements include capacity for multiplex sensing of biomolecules for higher accuracy, fast response time with short postprocessing procedures, a long life and cost effectiveness as to increase affordability of the biosensor. Existing biosensors are restricted due to their inability to achieve one or more of the mentioned requirements at the commercial stage and hence the ongoing research is focused on overcoming these limitations to make more accurate and sensitive sensors.

To be able to use graphene-based sensors for widespread commercial purposes, research efforts are focused at finding effective capturing agents to bind target molecules to the sensor surface and looking for solutions to the problem of the susceptibility of graphene's electrical conductivity to environmental changes, which hinder reliable and regular sensing process and results[125]. For biological applications, CNTs are functionalized with bioreceptors such as proteins or other microorganisms to result in CNT-bioFETs. While such devices offer significant advantages, one limitation is the background electrostatic noise, which can limit the performance of the sensor[126]. Some of the potential challenges when employing QDs as imaging probes include QDs with sizes less than 5nm being eliminated by renal filtration[127]. Additionally, toxicity studies of the nanoparticle cores used as sensor components could reveal the extent of their biocompatibility in biological systems [128, 129].

Notably, graphene and its aforementioned derivatives, beyond the well-researched applications in electronics, are noteworthy candidates for biomedical purposes as well, for instance in drug delivery[130], gene therapy[131], DNA sequencing[132], bio-imaging[133] and biosensing[134]. Considering the impact of biosensors in detection and management of infectious diseases, it will be worthwhile to overcome the current challenges related to their sensitivity, specificity, portability for point of care devices and cost upon commercialization. Current research in biosensors thus aims to develop robust biosensors capable of regeneration for reuse to allow use over a long time reduce cost issues. The key to developing better and more powerful biosensors seems to lie in the effective combination of polymers, nanomaterials and biology.

V. ACKNOWLEDGEMENTS

Dr. Vinod Singh is thankful to the Delhi Technological University for providing financial support under the scheme 'Research project grant to the faculty of DTU' (No. F. No. DTU/IRD/619/2019/2127). Dr. Vinod Singh is also thankful to AICTE for providing financial support under the scheme "Research Promotion Scheme" (Fife No. 8-123/FDC/RPS(POLICY-1)/2019-2020).

References

[1] R. Vargas-Bernal, "Electrical Properties of Two-Dimensional Materials Used in Gas Sensors," *Sensors*, vol. 19, no. 6, 2019, doi: 10.3390/s19061295.

[2] B. Jia, "2D optical materials and the implications for photonics," *APL Photonics*, vol. 4, no. 8, 2019, doi: 10.1063/1.5120030.

[3] H. Jiang, L. Zheng, Z. Liu, and X. Wang, "Two-dimensional materials: From mechanical properties to flexible mechanical sensors," *InfoMat*, 2019, doi: 10.1002/inf2.12072.

[4] Y.-Y. Chang, H. N. Han, and M. Kim, "Analyzing the microstructure and related properties of 2D materials by transmission electron microscopy," *Applied Microscopy*, vol. 49, no. 1, 2019, doi: 10.1186/s42649-019-0013-5.

[5] D. J. Late, A. Bhat, and C. S. Rout, "Fundamentals and Properties of 2D Materials in General and Sensing Applications," in *Fundamentals and Sensing Applications of 2D Materials*, 2019, pp. 5-24.

[6] H. Morgan, C. S. Rout, and D. J. Late, "Future Prospects of 2D Materials for Sensing Applications," in *Fundamentals and Sensing Applications of 2D Materials*, 2019, pp. 481-482.

[7] C. Mackin *et al.*, "Chemical sensor systems based on 2D and thin film materials," *2D Materials*, vol. 7, no. 2, 2020, doi: 10.1088/2053-1583/ab6e88.

[8] F. Yavari and N. Koratkar, "Graphene-Based Chemical Sensors," *The Journal of Physical Chemistry Letters*, vol. 3, no. 13, pp. 1746-1753, 2012/07/05 2012, doi: 10.1021/jz300358t.

[9] N. Atta, E. El-Ads, and A. Galal, "Graphene — A Platform for Sensor and Biosensor Applications," 2015, pp. 38-84.

[10] X. Kang, J. Wang, H. Wu, J. Liu, I. A. Aksay, and Y. Lin, "A graphene-based electrochemical sensor for sensitive detection of paracetamol," *Talanta*, vol. 81, no. 3, pp. 754-9, May 15 2010, doi: 10.1016/j.talanta.2010.01.009.

[11] F. Li, J. Li, Y. Feng, L. Yang, and Z. Du, "Electrochemical behavior of graphene doped carbon paste electrode and its application for sensitive determination of ascorbic acid," *Sensors and Actuators B: Chemical*, vol. 157, no. 1, pp. 110-114, 2011, doi: 10.1016/j.snb.2011.03.033.

[12] L. Chen, Y. Tang, K. Wang, C. Liu, and S. Luo, "Direct electrodeposition of reduced graphene oxide on glassy carbon electrode and its electrochemical application," *Electrochemistry Communications*, vol. 13, no. 2, pp. 133-137, 2011, doi: 10.1016/j.elecom.2010.11.033.

[13] T. Li, M. Yang, and H. Li, "Label-free electrochemical detection of cancer marker based on graphene–cobalt hexacyanoferrate nanocomposite," *Journal of Electroanalytical Chemistry*, vol. 655, no. 1, pp. 50-55, 2011, doi: 10.1016/j.jelechem.2011.02.009.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

[14] K. Liu, J. Zhang, G. Yang, C. Wang, and J.-J. Zhu, "Direct electrochemistry and electrocatalysis of hemoglobin based on poly(diallyldimethylammonium chloride) functionalized graphene sheets/room temperature ionic liquid composite film," *Electrochemistry Communications*, vol. 12, no. 3, pp. 402-405, 2010, doi: 10.1016/j.elecom.2010.01.004.

[15] W. Sun *et al.*, "Electrochemical sensor for transgenic maize MON810 sequence with electrostatic adsorption DNA on electrochemical reduced graphene modified electrode," *Sensors and Actuators B: Chemical*, vol. 202, pp. 160–166, 10/01 2014, doi: 10.1016/j.snb.2014.05.072.

[16] H. Wu *et al.*, "Glucose biosensor based on immobilization of glucose oxidase in platinum nanoparticles/graphene/chitosan nanocomposite film," *Talanta*, vol. 80, no. 1, pp. 403-6, Nov 15 2009, doi: 10.1016/j.talanta.2009.06.054.

[17] C. Anichini, W. Czepa, D. Pakulski, A. Aliprandi, A. Ciesielski, and P. Samorì, "Chemical sensing with 2D materials," *Chemical Society Reviews*, vol. 47, no. 13, pp. 4860-4908, 2018, doi: 10.1039/c8cs00417j.

[18] S. S. Sonakshi Saini, Maya Khangembam, Vinod Singh, "Structural analysis of the chemical vapour deposition grown molybdenum disulphide nanofilms for multifaceted applications.," *Indian Journal of Science and Technology.*, vol. 13, no. 29, pp. 2973-2980, 2020, doi: 10.17485/IJST/v13i29.1215.

[19] X. Wang, F. Nan, J. Zhao, T. Yang, T. Ge, and K. Jiao, "A label-free ultrasensitive electrochemical DNA sensor based on thin-layer MoS2 nanosheets with high electrochemical activity," *Biosens Bioelectron*, vol. 64, pp. 386-91, Feb 15 2015, doi: 10.1016/j.bios.2014.09.030.

[20] A. M. Parra-Alfambra, E. Casero, L. Vázquez, C. Quintana, M. del Pozo, and M. D. Petit-Domínguez, "MoS2 nanosheets for improving analytical performance of lactate biosensors," *Sensors and Actuators B: Chemical*, vol. 274, pp. 310-317, 2018, doi: 10.1016/j.snb.2018.07.124.

[21] W. Zhang, P. Zhang, Z. Su, and G. Wei, "Synthesis and sensor applications of MoS2-based nanocomposites," *Nanoscale*, vol. 7, no. 44, pp. 18364-18378, 2015, doi: 10.1039/c5nr06121k.

[22] J. Lee *et al.*, "Two-dimensional layered MoS(2) biosensors enable highly sensitive detection of biomolecules," *Sci Rep*, vol. 4, p. 7352, Dec 17 2014, doi: 10.1038/srep07352.

[23] K.-J. Huang, Y.-J. Liu, Y.-M. Liu, and L.-L. Wang, "Molybdenum disulfide nanoflower-chitosan-Au nanoparticles composites based electrochemical sensing platform for bisphenol A determination," *Journal of Hazardous Materials*, vol. 276, pp. 207-215, 2014, doi: 10.1016/j.jhazmat.2014.05.037.

[24] Y. Wang and J. T. W. Yeow, "A Review of Carbon Nanotubes-Based Gas Sensors," *Journal of Sensors*, vol. 2009, pp. 1-24, 2009, doi: 10.1155/2009/493904.

[25] I. V. Zaporotskova, N. P. Boroznina, Y. N. Parkhomenko, and L. V. Kozhitov, "Carbon nanotubes: Sensor properties. A review," *Modern Electronic Materials*, vol. 2, no. 4, pp. 95-105, 2016, doi: 10.1016/j.moem.2017.02.002.

[26] Y.-L. Zeng *et al.*, "Functionalization of multi-walled carbon nanotubes with poly(amidoamine) dendrimer for mediator-free glucose biosensor," *Electrochemistry Communications*, vol. 9, no. 1, pp. 185-190, 2007, doi: 10.1016/j.elecom.2006.08.052.

[27] N. Yang, X. Chen, T. Ren, P. Zhang, and D. Yang, "Carbon nanotube based biosensors," *Sensors and Actuators B: Chemical*, vol. 207, pp. 690-715, 2015, doi: 10.1016/j.snb.2014.10.040.

[28] R. A. Villamizar, A. Maroto, F. X. Rius, I. Inza, and M. J. Figueras, "Fast detection of Salmonella Infantis with carbon nanotube field effect transistors," *Biosens Bioelectron*, vol. 24, no. 2, pp. 279-83, Oct 15 2008, doi: 10.1016/j.bios.2008.03.046.

[29] K. J. Huang, Y. J. Liu, H. B. Wang, Y. Y. Wang, and Y. M. Liu, "Sub-femtomolar DNA detection based on layered molybdenum disulfide/multi-walled carbon nanotube composites, Au nanoparticle and enzyme multiple signal amplification," *Biosens Bioelectron*, vol. 55, pp. 195-202, May 15 2014, doi: 10.1016/j.bios.2013.11.061.

[30] V. Schroeder, S. Savagatrup, M. He, S. Lin, and T. M. Swager, "Carbon Nanotube Chemical Sensors," *Chemical Reviews*, vol. 119, no. 1, pp. 599-663, 2018, doi: 10.1021/acs.chemrev.8b00340.

[31] F. Patolsky, Gill, R., Weizmann, Y., Mokari, T., Banin, U., & Willner, I., "Lighting-Up the Dynamics of Telomerization and DNA Replication by CdSe−ZnS Quantum Dots," *Journal of the American Chemical Society*, no. 125, pp. 13918-13919, 2003, doi: DOI: 10.1021/ja035848c.

[32] M. Levy, S. F. Cater, and A. D. Ellington, "Quantum-dot aptamer beacons for the detection of proteins," *Chembiochem*, vol. 6, no. 12, pp. 2163-6, Dec 2005, doi: 10.1002/cbic.200500218.

[33] E. Chang *et al.*, "Protease-activated quantum dot probes," *Biochem Biophys Res Commun*, vol. 334, no. 4, pp. 1317-21, Sep 9 2005, doi: 10.1016/j.bbrc.2005.07.028.

[34] R. Freeman, T. Finder, R. Gill, and I. Willner, "Probing protein kinase (CK2) and alkaline phosphatase with CdSe/ZnS quantum dots," *Nano Lett*, vol. 10, no. 6, pp. 2192-6, Jun 9 2010, doi: 10.1021/nl101052f.

[35] R. Buiculescu, M. Hatzimarinaki, and N. A. Chaniotakis, "Biosilicated CdSe/ZnS quantum dots as photoluminescent transducers for acetylcholinesterase-based biosensors," *Anal Bioanal Chem*, vol. 398, no. 7-8, pp. 3015-21, Dec 2010, doi: 10.1007/s00216-010-4253-z.

[36] J. Liu, Lee, J. H., & Lu, Y., "Quantum Dot Encoding of Aptamer-LinkedNanostructures for One-Pot SimultaneousDetection of Multiple Analytes," *Analytical Chemistry*, vol. 11, no. 79, pp. 4120-4125, 2007, doi: DOI: 10.1021/ac070055k.

[37] P. K. Bae, K. N. Kim, S. J. Lee, H. J. Chang, C. K. Lee, and J. K. Park, "The modification of quantum dot probes used for the targeted imaging of his-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

tagged fusion proteins," *Biomaterials*, vol. 30, no. 5, pp. 836-42, Feb 2009, doi: 10.1016/j.biomaterials.2008.10.049.

[38] N. Chaniotakis and R. Buiculescu, "Semiconductor quantum dots in chemical sensors and biosensors," in *Nanosensors for Chemical and Biological Applications*, 2014, pp. 267-294.

[39] E. Hwang, "Bioluminescence resonance energy transfer (BRET) - based nanostructured biosensor for detection of glucose," The University of Western Ontario, Electronic Thesis and Dissertation Repository, 6782, 2019.

[40] M. Kumar, D. Zhang, D. Broyles, and S. K. Deo, "A rapid, sensitive, and selective bioluminescence resonance energy transfer (BRET)-based nucleic acid sensing system," *Biosens Bioelectron*, vol. 30, no. 1, pp. 133-9, Dec 15 2011, doi: 10.1016/j.bios.2011.08.043.

[41] E. Hwang, J. Song, and J. Zhang, "Integration of Nanomaterials and Bioluminescence Resonance Energy Transfer Techniques for Sensing Biomolecules," *Biosensors (Basel)*, vol. 9, no. 1, Mar 16 2019, doi: 10.3390/bios9010042.

[42] R. Wang, M. Guo, Y. Hu, J. Zhou, R. Wu, and X. Yang, "A Molecularly Imprinted Fluorescence Sensor Based on the ZnO Quantum Dot Core-Shell Structure for High Selectivity and Photolysis Function of Methylene Blue," *ACS Omega*, vol. 5, no. 32, pp. 20664-20673, Aug 18 2020, doi: 10.1021/acsomega.0c03095.

[43] Z. Song *et al.*, "Fully Stretchable and Humidity-Resistant Quantum Dot Gas Sensors," *ACS Sens*, vol. 3, no. 5, pp. 1048-1055, May 25 2018, doi: 10.1021/acssensors.8b00263.

[44] Y. Huang, W. Jiao, Z. Chu, X. Nie, R. Wang, and X. He, "SnS2 Quantum Dot-Based Optoelectronic Flexible Sensors for Ultrasensitive Detection of NO2 Down to 1 ppb," *ACS Appl Mater Interfaces*, vol. 12, no. 22, pp. 25178-25188, Jun 3 2020, doi: 10.1021/acsami.0c05240.

[45] H. Liu *et al.*, "Solution-Processed Gas Sensors Employing SnO2 Quantum Dot/MWCNT Nanocomposites," *ACS Appl Mater Interfaces*, vol. 8, no. 1, pp. 840-6, Jan 13 2016, doi: 10.1021/acsami.5b10188.

[46] M. Epifani *et al.*, "Solvothermal, chloroalkoxide-based synthesis of monoclinic WO(3) quantum dots and gas-sensing enhancement by surface oxygen vacancies," *ACS Appl Mater Interfaces*, vol. 6, no. 19, pp. 16808-16, Oct 8 2014, doi: 10.1021/am504158r.

[47] Z. Yue *et al.*, "Quantum-dot-based photoelectrochemical sensors for chemical and biological detection," *ACS Appl Mater Interfaces*, vol. 5, no. 8, pp. 2800-14, Apr 24 2013, doi: 10.1021/am3028662.

[48] V. Pardo-Yissar, Katz, E., Wasserman, J., & Willner, I., "Acetylcholine Esterase-Labeled CdS Nanoparticles on Electrodes: Photoelectrochemical Sensing of the Enzyme Inhibitors," *Journal of the American Chemical Society*, no. 125, pp. 622–623, 2003, doi: doi:10.1021/ja028922k.

[49] A. Shamirian, H. S. Afsari, A. Hassan, L. W. Miller, and P. T. Snee, "In vitro Detection of Hypoxia using a Ratiometric Quantum Dot-based Oxygen Sensor," *ACS Sens*, vol. 1, no. 10, pp. 1244-1250, Oct 28 2016, doi: 10.1021/acssensors.6b00452.

[50] A. S. Freires, F. M. dos Reis Lima, S. Yotsumoto-Neto, S. M. Silva, F. S. Damos, and R. de Cássia Silva Luz, "Exploiting CdSe/ZnS core-shell photocatalyst modified with cytochrome c for epinephrine determination in drugs utilized in cardiopulmonary resuscitation," *Microchemical Journal*, vol. 139, pp. 18-23, 2018, doi: 10.1016/j.microc.2018.02.008.

[51] R. Sharma, N. Kodan, V. Singh, S. N. Sharma, and O. P. Sinha, "Enhanced photoelectrochemical performance of TiO2 photoanode decorated with Pd-carbon core shell nanoparticles," *Renewable Energy*, vol. 134, pp. 1232-1239, 2019/04/01/ 2019, doi: https://doi.org/10.1016/j.renene.2018.09.061.

[52] V. Singh, S. Dhall, A. Kaushal, and B. R. Mehta, "Room temperature response and enhanced hydrogen sensing in size selected Pd-C core-shell nanoparticles: Role of carbon shell and Pd-C interface," *International Journal of Hydrogen Energy*, vol. 43, no. 2, pp. 1025-1033, 2018/01/11/ 2018, doi: https://doi.org/10.1016/j.ijhydene.2017.11.143.

[53] U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, and T. Nann, "Quantum dots versus organic dyes as fluorescent labels," *Nat Methods*, vol. 5, no. 9, pp. 763-75, Sep 2008, doi: 10.1038/nmeth.1248.

[54] K. Patir and S. K. Gogoi, "Nitrogen-doped carbon dots as fluorescence ON–OFF–ON sensor for parallel detection of copper(ii) and mercury(ii) ions in solutions as well as in filter paper-based microfluidic device," *Nanoscale Advances*, vol. 1, no. 2, pp. 592-601, 2019, doi: 10.1039/c8na00080h.

[55] U. Baruah, N. Gogoi, A. Konwar, M. Jyoti Deka, D. Chowdhury, and G. Majumdar, "Carbon Dot Based Sensing of Dopamine and Ascorbic Acid," *Journal of Nanoparticles*, vol. 2014, pp. 1-8, 2014, doi: 10.1155/2014/178518.

[56] Q. Liu *et al.*, "Distinguish cancer cells based on targeting turn-on fluorescence imaging by folate functionalized green emitting carbon dots," *Biosens Bioelectron*, vol. 64, pp. 119-25, Feb 15 2015, doi: 10.1016/j.bios.2014.08.052.

[57] C. I. Weng *et al.*, "One-step synthesis of biofunctional carbon quantum dots for bacterial labeling," *Biosens Bioelectron*, vol. 68, pp. 1-6, Jun 15 2015, doi: 10.1016/j.bios.2014.12.028.

[58] N. Wang, Y. Wang, T. Guo, T. Yang, M. Chen, and J. Wang, "Green preparation of carbon dots with papaya as carbon source for effective fluorescent sensing of Iron (III) and Escherichia coli," *Biosens*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

Bioelectron, vol. 85, pp. 68-75, Nov 15 2016, doi: 10.1016/j.bios.2016.04.089.

[59] M. Louleb *et al.*, "Detection of Dopamine in Human Fluids Using N-Doped Carbon Dots," *ACS Applied Nano Materials*, vol. 3, no. 8, pp. 8004-8011, 2020, doi: 10.1021/acsanm.0c01461.

[60] P. Li *et al.*, "Group IV nanodots: Newly emerging properties and application in biomarkers sensing," *TrAC Trends in Analytical Chemistry*, vol. 131, 2020, doi: 10.1016/j.trac.2020.116007.

[61] Y. Xu, X. Wang, W. L. Zhang, F. Lv, and S. Guo, "Recent progress in two-dimensional inorganic quantum dots," *Chem Soc Rev*, vol. 47, no. 2, pp. 586-625, Jan 22 2018, doi: 10.1039/c7cs00500h.

[62] A. A. Ensafi, P. Nasr-Esfahani, and B. Rezaei, "Metronidazole determination with an extremely sensitive and selective electrochemical sensor based on graphene nanoplatelets and molecularly imprinted polymers on graphene quantum dots," *Sensors and Actuators B: Chemical*, vol. 270, pp. 192-199, 2018, doi: 10.1016/j.snb.2018.05.024.

[63] Q. Huang, X. Lin, L. Tong, and Q.-X. Tong, "Graphene Quantum Dots/Multiwalled Carbon Nanotubes Composite-Based Electrochemical Sensor for Detecting Dopamine Release from Living Cells," *ACS Sustainable Chemistry & Engineering*, vol. 8, no. 3, pp. 1644-1650, 2020, doi: 10.1021/acssuschemeng.9b06623.

[64] Y. Zhang *et al.*, "Graphene quantum dots/gold electrode and its application in living cell H2O2 detection," *Nanoscale*, vol. 5, no. 5, pp. 1816-9, Mar 7 2013, doi: 10.1039/c3nr33954h.

[65] CDC, "Identifying the source of the outbreak," Jul. 01, 2020. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/about-epidemiology/identifying-source-outbreak.html (accessed Oct. 24, 2020).

[66] V. C. C. Cheng, S. K. P. Lau, P. C. Y. Woo, and K. Y. Yuen, "Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection," (in eng), *Clin Microbiol Rev*, vol. 20, no. 4, pp. 660-694, 2007, doi: 10.1128/CMR.00023-07.

[67] Y.-W. Tang, J. E. Schmitz, D. H. Persing, and C. W. Stratton, "Laboratory Diagnosis of COVID-19: Current Issues and Challenges," *Journal of Clinical Microbiology*, vol. 58, no. 6, pp. e00512-20, 2020, doi: 10.1128/jcm.00512-20.

[68] R. Kumar, S. Nagpal, S. Kaushik, and S. Mendiratta, "COVID-19 diagnostic approaches: different roads to the same destination," *VirusDisease*, vol. 31, no. 2, pp. 97-105, 2020/06/01 2020, doi: 10.1007/s13337-020-00599-7.

[69] H. A. Hussein, R. Y. A. Hassan, M. Chino, and F. Febbraio, "Point-of-Care Diagnostics of COVID-19: From Current Work to Future Perspectives," *Sensors*, vol. 20, no. 15, p. 4289, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/15/4289.

[70] J. Chen *et al.*, "COVID-19 infection: the China and Italy perspectives," *Cell Death & Disease*, vol. 11, no. 6, p. 438, 2020/06/08 2020, doi: 10.1038/s41419-020-2603-0.

[71] M. Chung *et al.*, "CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202-207, 2020, doi: 10.1148/radiol.2020200230.

[72] "What tests could potentially be used for the screening, diagnosis and monitoring of COVID-19 and what are their advantages and disadvantages? - CEBM." https://www.cebm.net/covid-19/what-tests-could-potentially-be-used-for-the-screening-diagnosis-and-monitoring-of-covid-19-and-what-are-their-advantages-and-disadvantages/ (accessed Oct. 24, 2020).

[73] Y. Fang et al., "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," Radiology, vol. 296, no. 2, pp. E115–E117, Aug. 2020, doi: 10.1148/radiol.2020200432.

[74] T. Ai et al., "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," Radiology, vol. 296, no. 2, pp. E32–E40, Aug. 2020, doi: 10.1148/radiol.2020200642.

[75] "PCR - FosunPharmaUSA." http://fosunpharmausa.com/covid19/pcr/ (accessed Oct. 28, 2020).

[76] "Smart Detect SARS-CoV-2 rRT-PCR Kit/48T." https://supply.unicef.org/s0003891.html (accessed Oct. 28, 2020).

[77] "FDA Emergency Use Authorized COVID-19 IgG/IgM Rapid Test Cassette." https://www.confirmbiosciences.com/covid19-instant-coronavirus-test-kit/ (accessed Oct. 28, 2020).

[78]L. R. Feldstein et al., "Multisystem Inflammatory Syndrome in U.S. Children and Adolescents," N. Engl. J. Med., vol. 383, no. 4, pp. 334–346, Jul. 2020, doi: 10.1056/NEJMoa2021680.

[79] K. M. Koczula and A. Gallotta, "Lateral flow assays," Essays Biochem., vol. 60, no. 1, pp. 111–120, Jun. 2016, doi: 10.1042/EBC20150012.

[80] Y. Pan et al., "Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China," Eur. Radiol., vol. 30, no. 6, pp. 3306–3309, Jun. 2020, doi: 10.1007/s00330-020-06731-x.

[81] X. Xu et al., "Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2," Eur. J. Nucl. Med. Mol. Imaging, vol. 47, no. 5, pp. 1275–1280, May 2020, doi: 10.1007/s00259-020-04735-9.

[82] J. Lei, J. Li, X. Li, and X. Qi, "CT Imaging of the 2019 Novel Coronavirus (2019-nCoV) Pneumonia," Radiology, vol. 295, no. 1, p. 18, Apr. 2020, doi: 10.1148/radiol.2020200236.

[83] Y. Fang et al., "Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR," Radiology, vol. 296, no. 2, pp. E115–E117, Aug. 2020, doi: 10.1148/radiol.2020200432.

[84] T. Ai et al., "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," Radiology, vol.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

296, no. 2, pp. E32–E40, Aug. 2020, doi: 10.1148/radiol.2020200642.

[85] H. X. Bai et al., "Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT," Radiology, vol. 296, no. 2, pp. E46–E54, Aug. 2020, doi: 10.1148/radiol.2020200823.

[86] Y.-H. Yun et al., "Tiny Medicine: Nanomaterial-Based Biosensors," Sensors, vol. 9, no. 11. pp. 9275–9299, 2009, doi: 10.3390/s91109275.

[87] N. Bhalla, Y. Pan, Z. Yang, and A. F. Payam, "Opportunities and Challenges for Biosensors and Nanoscale Analytical Tools for Pandemics: COVID-19," ACS Nano, vol. 14, no. 7, pp. 7783–7807, Jul. 2020, doi: 10.1021/acsnano.0c04421.

[88] M. Holzinger, A. Le Goff, and S. Cosnier, "Nanomaterials for biosensing applications: a review," Frontiers in Chemistry, vol. 2. 2014, doi: 10.3389/fchem.2014.00063.

[89] M. Qiu *et al.*, "Antibody responses to individual proteins of SARS coronavirus and their neutralization activities," *Microbes and Infection*, vol. 7, no. 5, pp. 882-889, 2005/05/01/ 2005, doi: https://doi.org/10.1016/j.micinf.2005.02.006.

[90] Y.-J. Tan *et al.*, "Profiles of Antibody Responses against Severe Acute Respiratory Syndrome Coronavirus Recombinant Proteins and Their Potential Use as Diagnostic Markers," *Clinical and Diagnostic Laboratory Immunology*, vol. 11, no. 2, pp. 362-371, 2004, doi: 10.1128/cdli.11.2.362-371.2004.

[91] L. Lu *et al.*, "Immunological Characterization of the Spike Protein of the Severe Acute Respiratory Syndrome Coronavirus," *Journal of Clinical Microbiology*, vol. 42, no. 4, pp. 1570-1576, 2004, doi: 10.1128/jcm.42.4.1570-1576.2004.

[92] S. Mavrikou, G. Moschopoulou, V. Tsekouras, and S. Kintzios, "Development of a Portable, Ultra-Rapid and Ultra-Sensitive Cell-Based Biosensor for the Direct Detection of the SARS-CoV-2 S1 Spike Protein Antigen," *Sensors*, vol. 20, no. 11, p. 3121, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/11/3121.

[93] R. J. G. Hulswit *et al.*, "Human coronaviruses OC43 and HKU1 bind to 9-O-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A," *Proceedings of the National Academy of Sciences*, vol. 116, no. 7, pp. 2681-2690, 2019, doi: 10.1073/pnas.1809667116.

[94] Y.-J. Park *et al.*, "Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors," *Nature Structural & Molecular Biology*, vol. 26, no. 12, pp. 1151-1157, 2019/12/01 2019, doi: 10.1038/s41594-019-0334-7.

[95] D. Murugan, H. Bhatia, V. V. R. Sai, and J. Satija, "P-FAB: A Fiber-Optic Biosensor Device for Rapid Detection of COVID-19," *Transactions of the Indian National Academy of Engineering*, vol. 5, no. 2, pp. 211-215, 2020/06/01 2020, doi: 10.1007/s41403-020-00122-w.

[96] E. Williams, K. Bond, B. Zhang, M. Putland, and D. A. Williamson, "Saliva as a Noninvasive Specimen for Detection of SARS-CoV-2," *Journal of Clinical Microbiology*, vol. 58, no. 8, pp. e00776-20, 2020, doi: 10.1128/jcm.00776-20.

[97] W. Liu *et al.*, "Evaluation of Nucleocapsid and Spike Protein-Based Enzyme-Linked Immunosorbent Assays for Detecting Antibodies against SARS-CoV-2," *Journal of Clinical Microbiology*, vol. 58, no. 6, pp. e00461-20, 2020, doi: 10.1128/jcm.00461-20.

[98] A. L. Wyllie *et al.*, "Saliva is more sensitive for SARS-CoV-2 detection in COVID-19 patients than nasopharyngeal swabs," *medRxiv*, p. 2020.04.16.20067835, 2020, doi: 10.1101/2020.04.16.20067835.

[99] R. K. Verma and B. D. Gupta, "Theoretical modelling of a bi-dimensional U-shaped surface plasmon resonance based fibre optic sensor for sensitivity enhancement," *Journal of Physics D: Applied Physics*, vol. 41, no. 9, p. 095106, 2008/04/07 2008, doi: 10.1088/0022-3727/41/9/095106.

[100] S. F. L. f. M. S. a. T. (EMPA). "A new biosensor for the COVID-19 virus: Detection in the environment." www.sciencedaily.com/releases/2020/04/200421112520.htm (accessed August 30, 2020).

[101] G. Qiu, Z. Gai, Y. Tao, J. Schmitt, G. A. Kullak-Ublick, and J. Wang, "Dual-Functional Plasmonic Photothermal Biosensors for Highly Accurate Severe Acute Respiratory Syndrome Coronavirus 2 Detection," *ACS Nano*, vol. 14, no. 5, pp. 5268-5277, 2020/05/26 2020, doi: 10.1021/acsnano.0c02439.

[102] J. R. Mejía-Salazar and O. N. Oliveira, "Plasmonic Biosensing," *Chemical Reviews*, vol. 118, no. 20, pp. 10617-10625, 2018/10/24 2018, doi: 10.1021/acs.chemrev.8b00359.

[103] S. Firdous, S. Anwar, and R. Rafya, "Development of surface plasmon resonance (SPR) biosensors for use in the diagnostics of malignant and infectious diseases," *Laser Physics Letters*, vol. 15, no. 6, p. 065602, 2018/04/26 2018, doi: 10.1088/1612-202x/aab43f.

[104] M. Soler, C. S. Huertas, and L. M. Lechuga, "Label-free plasmonic biosensors for point-of-care diagnostics: a review," *Expert Review of Molecular Diagnostics*, vol. 19, no. 1, pp. 71-81, 2019/01/02 2019, doi: 10.1080/14737159.2019.1554435.

[105] J.-F. Masson, "Surface Plasmon Resonance Clinical Biosensors for Medical Diagnostics," *ACS Sensors*, vol. 2, no. 1, pp. 16-30, 2017/01/27 2017, doi: 10.1021/acssensors.6b00763.

[106] J. Liu, M. Jalali, S. Mahshid, and S. Wachsmann-Hogiu, "Are plasmonic optical biosensors ready for use in point-of-need applications?," *Analyst*, 10.1039/C9AN02149C vol. 145, no. 2, pp. 364-384, 2020, doi: 10.1039/C9AN02149C.

[107] X. Michalet *et al.*, "Quantum Dots for Live Cells, in Vivo Imaging, and Diagnostics," *Science*, vol. 307, no. 5709, pp. 538-544, 2005, doi: 10.1126/science.1104274.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

[108] S. R. Mudshinge, A. B. Deore, S. Patil, and C. M. Bhalgat, "Nanoparticles: Emerging carriers for drug delivery," *Saudi Pharmaceutical Journal*, vol. 19, no. 3, pp. 129-141, 2011/07/01/ 2011, doi: https://doi.org/10.1016/j.jsps.2011.04.001.

[109] M. F. Frasco and N. Chaniotakis, "Semiconductor Quantum Dots in Chemical Sensors and Biosensors," *Sensors*, vol. 9, no. 9, pp. 7266-7286, 2009. [Online]. Available: https://www.mdpi.com/1424-8220/9/9/7266.

[110] B. A. Makwana, S. Darjee, V. K. Jain, A. Kongor, G. Sindhav, and M. V. Rao, "A comparative study: Metal nanoparticles as fluorescent sensors for biomolecules and their biomedical application," *Sensors and Actuators B: Chemical*, vol. 246, pp. 686-695, 2017/07/01/ 2017, doi: https://doi.org/10.1016/j.snb.2017.02.054.

[111] B. Eftekhari-Sis, S. Karaminejad, F. Malekan, H. Y. Araghi, and A. Akbari, "CdSe Quantum Dots Based Nano-Biosensor for Detection of 185delAG Mutation in BRCA1 Gene, Responsible for Breast Cancer," *Journal of Inorganic and Organometallic Polymers and Materials*, vol. 27, no. 6, pp. 1911-1917, 2017/11/01 2017, doi: 10.1007/s10904-017-0661-2.

[112] N. Hoa *et al.*, "Fluorescence biosensor based on CdTe quantum dots for specific detection of H5N1 avian influenza virus," *Adv. Nat. Sci.: Nanosci. Nanotechnol*, vol. 3, 09/01 2012, doi: 10.1088/2043-6262/3/3/035014.

[113] R. D. Singh *et al.*, "Quantum Dot Based Nano-Biosensors for Detection of Circulating Cell Free miRNAs in Lung Carcinogenesis: From Biology to Clinical Translation," (in English), *Frontiers in Genetics*, Review vol. 9, no. 616, 2018-December-06 2018, doi: 10.3389/fgene.2018.00616.

[114] T. Pasinszki, M. Krebsz, T. T. Tung, and D. Losic, "Carbon Nanomaterial Based Biosensors for Non-Invasive Detection of Cancer and Disease Biomarkers for Clinical Diagnosis," *Sensors*, vol. 17, no. 8, p. 1919, 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/8/1919.

[115] D. Bhattacharyya, P. K. Sarswat, and M. L. Free, "Quantum dots and carbon dots based fluorescent sensors for TB biomarkers detection," *Vacuum*, vol. 146, pp. 606-613, 2017/12/01/ 2017, doi: https://doi.org/10.1016/j.vacuum.2017.02.003.

[116] Y. Li, H. J. Schluesener, and S. Xu, "Gold nanoparticle-based biosensors," *Gold Bulletin*, vol. 43, no. 1, pp. 29-41, 2010/03/01 2010, doi: 10.1007/BF03214964.

[117] H. C. Lai, S. F. Chin, S. C. Pang, M. S. Henry Sum, and D. Perera, "Carbon Nanoparticles Based Electrochemical Biosensor Strip for Detection of Japanese Encephalitis Virus," *Journal of Nanomaterials*, vol. 2017, p. 3615707, 2017/12/18 2017, doi: 10.1155/2017/3615707.

[118] N. Sanvicens, C. Pastells, N. Pascual, and M. P. Marco, "Nanoparticle-based biosensors for detection of pathogenic bacteria," *TrAC Trends in Analytical Chemistry*, vol. 28, no. 11, pp. 1243-1252, 2009/12/01/ 2009, doi: https://doi.org/10.1016/j.trac.2009.08.002.

[119] K.-I. Joo *et al.*, "Site-Specific Labeling of Enveloped Viruses with Quantum Dots for Single Virus Tracking," *ACS Nano*, vol. 2, no. 8, pp. 1553-1562, 2008/08/26 2008, doi: 10.1021/nn8002136.

[120] M. Islam and M. Z. Ahsan, "Plausible Approach for Rapid Detection of SARS-CoV-2 Virus by Magnetic Nanoparticle Based Biosensors," vol. 6, pp. 6-13, 07/23 2020, doi: 10.11648/j.ajn.20200602.11.

[121] G. Seo *et al.*, "Rapid Detection of COVID-19 Causative Virus (SARS-CoV-2) in Human Nasopharyngeal Swab Specimens Using Field-Effect Transistor-Based Biosensor," *ACS Nano*, vol. 14, no. 4, pp. 5135-5142, 2020/04/28 2020, doi: 10.1021/acsnano.0c02823.

[122] J. Saleh, C. Peyssonnaux, K. K. Singh, and M. Edeas, "Mitochondria and microbiota dysfunction in COVID-19 pathogenesis," *Mitochondrion*, vol. 54, pp. 1-7, 2020/09/01/ 2020, doi: https://doi.org/10.1016/j.mito.2020.06.008.

[123] Z. S. Miripour *et al.*, "Real-time diagnosis of reactive oxygen species (ROS) in fresh sputum by electrochemical tracing; correlation between COVID-19 and viral-induced ROS in lung/respiratory epithelium during this pandemic," (in eng), *Biosensors & bioelectronics*, vol. 165, pp. 112435-112435, 2020, doi: 10.1016/j.bios.2020.112435.

[124] T. L. Freeman and T. H. Swartz, "Targeting the NLRP3 Inflammasome in Severe COVID-19," (in eng), *Front Immunol*, vol. 11, pp. 1518-1518, 2020, doi: 10.3389/fimmu.2020.01518.

[125] R. Malik, V. K. Tomer, and V. Chaudhary, "Hybridized Graphene for Chemical Sensing," in *Functionalized Graphene Nanocomposites and their Derivatives*, 2019, pp. 323-338.

[126] L. Camilli and M. Passacantando, "Advances on Sensors Based on Carbon Nanotubes," *Chemosensors*, vol. 6, no. 4, 2018, doi: 10.3390/chemosensors6040062.

[127] H. Soo Choi *et al.*, "Renal clearance of quantum dots," *Nature Biotechnology*, vol. 25, no. 10, pp. 1165-1170, 2007/10/01 2007, doi: 10.1038/nbt1340.

[128] M. Bottrill and M. Green, "Some aspects of quantum dot toxicity," *Chemical Communications*, 10.1039/C1CC10692A vol. 47, no. 25, pp. 7039-7050, 2011, doi: 10.1039/C1CC10692A.

[129] M. J. D. Clift and V. Stone, "Quantum Dots: An Insight and Perspective of Their Biological Interaction and How This Relates to Their Relevance for Clinical Use," *Theranostics*, Review vol. 2, no. 7, pp. 668-680, 2012, doi: 10.7150/thno.4545.

[130] Z. Liu, J. T. Robinson, X. Sun, and H. Dai, "PEGylated nanographene oxide for delivery of water-insoluble cancer drugs," (in eng), *J Am Chem Soc*, vol. 130, no. 33, pp. 10876-7, Aug 20 2008, doi: 10.1021/ja803688x.

[131] L. Feng, S. Zhang, and Z. Liu, "Graphene based gene transfection," (in eng), *Nanoscale*, vol. 3, no. 3, pp. 1252-7, Mar 2011, doi: 10.1039/c0nr00680g.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2020.3036748, IEEE Sensors Journal

[132] S. J. Heerema and C. Dekker, "Graphene nanodevices for DNA sequencing," *Nature Nanotechnology*, vol. 11, no. 2, pp. 127-136, 2016/02/01 2016, doi: 10.1038/nnano.2015.307.

[133] C. Peng, W. Hu, Y. Zhou, C. Fan, and Q. Huang, "Intracellular Imaging with a Graphene-Based Fluorescent Probe," *Small*, vol. 6, no. 15, pp. 1686-1692, 2010, doi: 10.1002/smll.201000560.

[134] Q. Zeng *et al.*, "Self-Assembled Graphene–Enzyme Hierarchical Nanostructures for Electrochemical Biosensing," *Advanced Functional Materials*, vol. 20, no. 19, pp. 3366-3372, 2010, doi: 10.1002/adfm.201000540.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 09:25:14 UTC from IEEE Xplore. Restrictions apply.

AIP Conference Proceedings

Numerical simulation of n-TiO2/p-CIGS solar cell

Cite as: AIP Conference Proceedings **2265**, 030675 (2020); https://doi.org/10.1063/5.0017203
Published Online: 05 November 2020

Gagandeep, Mukhtiyar Singh, Ramesh Kumar, and Sonia Kumari

View Online Export Citation

Meet the Next Generation of Quantum Analyzers
And Join the Launch Event on November 17th

Register now
Zurich Instruments

AIP Publishing

© 2020 Author(s).

Numerical Simulation of n-TiO2/p-CIGS Solar Cell

Gagandeep[1, a)], Mukhtiyar Singh[2], Ramesh Kumar[1, b)] and Sonia Kumari[1]

[1]*Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar, India*
[2]*Department of Applied Physics, Delhi Technological University, Delhi, India*

[a)]gagankundu22@gmail.com
Corresponding author: [b)] rameshkumar.bibiyan@gmail.com

Abstract. The mostly used buffer layer in CIGS solar cell is CdS. This buffer layer is toxic. Therefore, the p-type Cu (In, Ga)Se_2 (CIGS) absorber layer is simulated with non toxic and wide band gap TiO_2 material as n-type buffer layer in AFORS-HET software under the illumination condition of AM 1.5 G. To see the effect of absorber layer thickness on the performance parameters of the cell, it is varied from 200-1900 nm. The highest efficiency of the cell is obtained 16.03% for optimized thickness of absorber layer i.e., at 1900 nm. The temperature influences the performance of the cell and its performance is maximum at temperature 300 K. The external quantum efficiency behavior with wavelength is observed. The reported results in this study show that TiO_2 material is better replacement of CdS material.

Key Words: Solar Cell, Efficiency, TiO2, AFORS-HET, CIGS

INTRODUCTION

The chalcopyrite Cu(In, Ga)Se_2 (CIGS) material based solar cells is very promising thin film solar cell and play very important role in the enhancement of photovoltaic technology [1-3]. The CIGS solar cell has excellent outdoor stability. The CIGS based solar cells are attracting considerable space application interest because of its stability against particle irradiation. Its stability is proved by electron and proton irradiation test of CdTe and CIGS solar cells which is superior to Ⅲ-Ⅴ or silicon solar cells. The laboratory scale power conversion efficiency (PCE) of this cell is attaining 20.4% on a flexible polymer substrate [4] and 22.6% on a soda lime glass (SLG) substrate [5]. The CdS plays the role of buffer layer in CIGS solar cell and resulting the highest above mentioned device efficiency. The CdS layer on CIGS is formed by chemical bath deposition (CBD) method. The CdS layer have relatively low band gap energy (2.4-2.5 eV). Hence, the finest performance of the cell is limited by the parasitic absorption in short wavelength region [6]. The CdS buffer layer is also toxic in nature.

Therefore, different buffer layer with a wider band gap is proposed [7- 9]. W. Hsu et. al [10] experimentally studied that amorphous TiO_2 act as buffer layer on CIGS cell and improved the performance of the device

Keeping the observation in mind, we simulated the TiO_2 as buffer layer in CIGS solar cell and obtained the PCE is 16.03%. Our theoretical results reported herein are improved the performance of the device.

SIMULATION DETAILS

Here, we simulate a polycrystalline CIGS based solar cell by n-TiO_2 /p-CIGS heterojunction structure with ITO and Mo are front and back contact respectively. In this paper, we use AFORS-HET (Automat for simulation of heterostructure) [11-13] to analyze the working of CIGS based thin film solar cell. The simulated structure is shown in figure 1.

DAE Solid State Physics Symposium 2019
AIP Conf. Proc. 2265, 030675-1–030675-4; https://doi.org/10.1063/5.0017203
Published by AIP Publishing. 978-0-7354-2025-0/$30.00

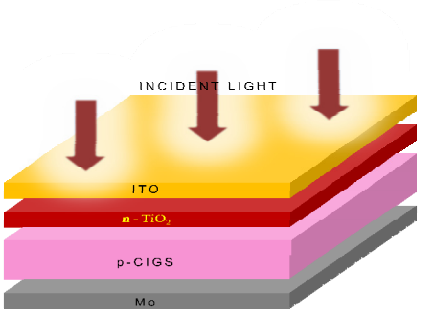


FIGURE 1. Schematic structure of simulated CIGS solar cell

The thickness of the CIGS layer is varied from 200-1900 nm. The variations in the values of the performance parameters of the cell such as short current density (J_{sc}), open circuit voltage (V_{oc}), fill factor (FF) and PCE (η) are noticed with absorber layer thickness. The impact of temperature on the performance parameters of the device is observed. The J-V (current density-voltage) characteristics are investigated for this structure with optimized absorber layer thickness. The quantum efficiency spectra with wavelength are also plotted. The layer parameters used in this work are taken from different theoretical and experimental studies [14-15] and mentioned in Table 1.

TABLE1. Layer parameters used in simulation

Parameter	**n-TiO_2**	**p-CIGS**
Layer thickness (nm)	90	200-1900
Dielectric constant	100	10
Electron affinity (eV)	4	4.2
Band gap (eV)	3.2	1.4
Optical band gap (eV)	3.2	1.4
Valence band density (cm^{-3})	1×10^{21}	2.2×10^{18}
Conduction band density (cm^{-3})	2×10^{20}	1.8×10^{19}
Hole mobility ($cm^2\ V^{-1}\ s^{-1}$)	0.006	100
Electron mobility ($cm^2\ V^{-1}\ s^{-1}$)	0.006	25
Thermal velocity of electron(cm/s)	1×10^{6}	1×10^{7}
Thermal velocity of hole(cm/s)	4×10^{6}	1×10^{7}
Doping concentration(cm^{-3})	5×10^{19}	2×10^{13}
Layer density(g cm^{-3})	2.328	2.328

RESULTS AND DISCUSSIONS

The thickness of CIGS layer is varied from 200-1900 nm and performance of the cell is observed. The cell parameters variation with absorber layer thickness is displayed in fig.2. The value of cell parameters such as J_{sc}, V_{oc}, FF and η are changed from 23.06 mA/cm^2 to 28.01 mA/cm^2, 606.4 mV to 697.7 mV, 81.95% to 82% and 11.46% to 16.03% respectively. It is explained based on the fact that larger thickness of CIGS layer absorbs more radiation and produce more carrier i.e. electron-hole pair and probably increase the performance parameters i.e. V_{oc}, J_{sc}, and η of the device. The FF of the cell is increased up to 82.24% at 1200 nm absorber thickness but decreases up to 82% at 1900 nm absorber thickness. The optimum thickness of CIGS layer is around 1900 nm from which the highest

efficiency of 16.03% is obtained for the n-TiO_2/p-CIGS cell. The results reported herein are accorded with the general trends of the device i.e. increased layer thickness improves the efficiency of the cell [16].

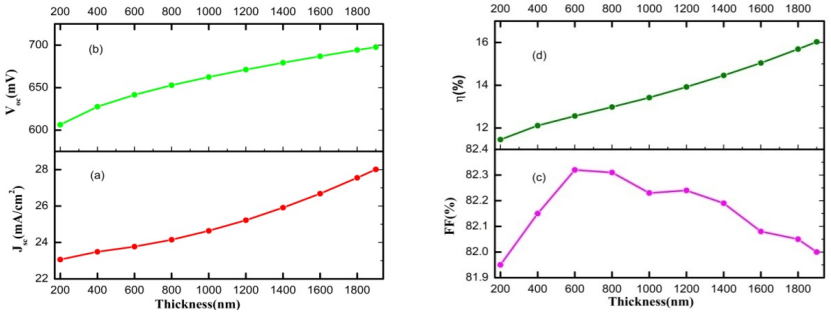


FIGURE2. Variation of solar cell parameters (a) J_{sc}, (b) V_{oc}, (c) FF and (d) Efficiency with absorber layer thickness

The variation of efficiency, FF, J_{sc} and V_{oc} with operating temperature is shown in figure 3. The performance of the cell, are also influenced by the operating temperature. The cell parameters variation is discussed in the temperature range 300-600 K under a constant illumination of 1000 W/m^2. The characteristics parameters of the cell namely V_{oc}, FF and η are strongly influenced by the operating temperature and get reduced with increase in temperature of the cell but J_{sc} remains constant. From the observed results we analyzed that the η of the cell is decreased from 16.03% to 0.15% in the range 300-600 K. It may be probably happened due to that, as we increase the temperature of the cell then there is increased in the reverse saturation current density J_o which is caused by the increase in the intrinsic carrier concentration and decrease in the band gap of the material. The conductivity of the cell is also affected by the temperature. At optimized value of absorber layer thickness J-V characteristics of the cell is also studied and obtain the η of 16.03% with FF of 82%. We also calculate the V_{oc} = 697.7 mV and J_{sc} = 28.01 mA/cm^2 respectively and results are shown in fig. 4(a).

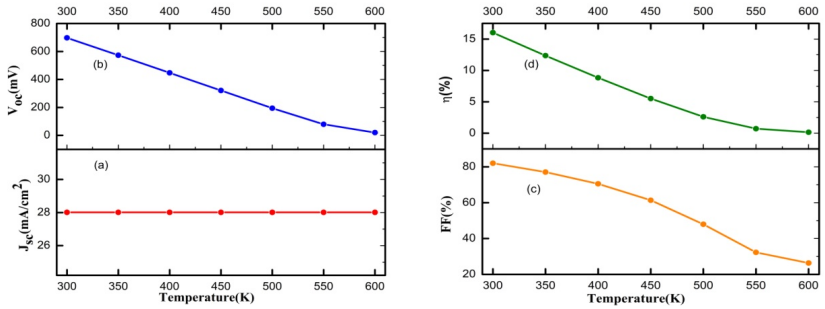


FIGURE3. Variation of solar cell parameters (a) J_{sc}, (b) V_{oc}, (c) FF and (d) Efficiency with operating temperature

The radiation absorbed by the cell is given by the spectra of EQE (external quantum efficiency). The QE spectra provides an information about the absorption of light by the CIGS device as compared to incident one. The EQE provides the detailed study about all photons whether they are absorbed, reflected or transmitted and provide us a maximum value in the wavelength range of 600-900 nm. All the absorbed photon by the cell are counted by the IQE (internal quantum efficiency) and gives maximum in 400-900 nm of wavelength but IQE1 becomes almost unity in this range of wavelength and counts only that photons which are absorbed by the electrical layers. The results of quantum efficiency are shown in fig 4 (b)

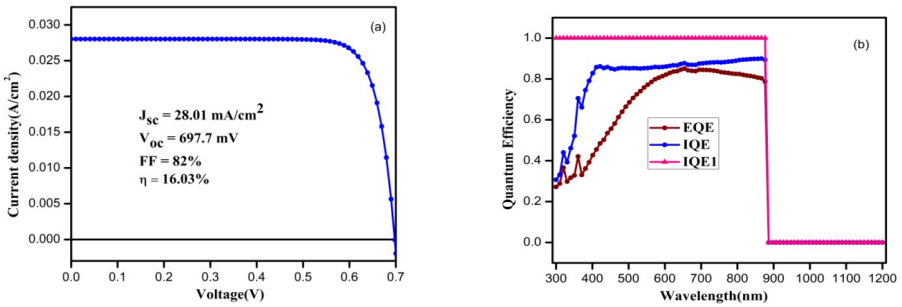


FIGURE4. (a) J-V characteristics (b) Quantum efficiency with wavelength

CONCLUSIONS

Theoretically, we simulated the TiO_2 layer as buffer layer in CIGS solar cell. The thickness of the absorber layer influences the performance of the device. The maximum value of power conversion efficiency is 16.03% at 1900 nm absorber thickness. The cell performance is influenced by the temperature and have attained maximum value at 300 K. The J-V curve and variation of external quantum efficiency of the solar cell with wavelength are also plotted. The obtained results in this study boost the performance of TiO_2 as buffer layer.

ACKNOWLEDGMENT

One of the authors (**Ramesh Kumar**) is thankful to the **University Grants Commission, INDIA** for giving the financial support through START-UP GRANT under FRPS (No. F.30-364/2017 (BSR), Dated 15/12/2017). The authors highly acknowledge the AFORS-HET software developer for providing free access of the software.

REFERENCES

1. J. Löckinger, S. Nishiwaki, T. P. Weiss, B. Bissig et al., Solar Energy Materials and Solar Cells, 174, 397 (2018).
2. C. Zhang, H. Zhu, X. Liang, D. Zhou, Y. Guo, X. Niu, Z. Li, J. Chen, Y. Mai, Vacuum, 141, 89 (2017).
3. Gagandeep, M. Singh and R. Kumar, AIP Conference Proceedings, 2093, 020018 (2019).
4. A. Chirila et.al., Nat. Mater. 12, 1107 (2013).
5. P. Jackson et.al., Phys. Status Solidi-R 10, 583(2016).
6. N. Naghavi et.al., Prog. Photovolt.18,411 (2010).
7. S. Spiering et.al.,Solar Energ.Mater. Sol.C 144, 544 (2016).
8. J. Lindahl et. al., IEEE J. Photovolt.3, 1100 (2013).
9. T. M.Friedlmeir et. al., IEEE J. Photovolt 5, 1487(2015).
10. W. Hsu et.al., Sci. Rep. 5, 16028 (2015).
11. R.Varache et. al., Sol. Energ. Mat. Sol. C.141, 14 (2015).
12. R. Stangl, J. Haschke and C. Leendertz,Solar Energy, Intech (R.D. Rugescu (Ed.), Croatia, 2009), pp. 432.
13. R.Stangl, M. Kriegel and M. Schmidt, Proceeding of the WCPEC-4, Fourth World Conference on Photovoltaic Energy Conversion (Hawaii, USA), PP. 1350 (2006).
14. M. Mostefaoui et. al., Energy Procedia, 74, 736 (2015).
15. S. Dabbabi, T. B. Nasr, N. K. Turki, Results in Physics, 7, 4020 (2017).
16. S. Tobbeche and H. Amar, JNTM, 04, 89 (2014).

arXiv:2011.11404v1 [math.CV] 23 Nov 2020

ON CONVEX DOMINANTS OF EXACT DIFFERENTIAL SUBORDINATION

S. SIVAPRASAD KUMAR AND SHAGUN BANGA

Abstract. Let h be a non vanishing convex univalent function and p be an analytic function in $\mathbb{D}$. We consider the differential subordination

$$\psi_i(p(z), zp'(z)) \prec h(z)$$

with the admissible functions in consideration as $\psi_1 := (\beta p(z)+\gamma)^{-\alpha}\left(\frac{(\beta p(z)+\gamma)}{\beta(1-\alpha)} + zp'(z)\right)$ and $\psi_2 := \frac{1}{\sqrt{\gamma\beta}} \arctan\left(\sqrt{\frac{\beta}{\gamma}} p^{1-\alpha}(z)\right) + \left(\frac{1-\alpha}{\beta p^{2(1-\alpha)}(z)+\gamma}\right)\frac{zp'(z)}{p^{\alpha}(z)}$. The objective of this paper is to find the dominants, preferably the best dominant(say q) of the solution of the above differential subordination satisfying $\psi_i(q, nzq'(z)) = h(z)$. Further, we show that $\psi_i(q, zq'(z)) = h(z)$ is an exact differential equation and q is a convex univalent function in $\mathbb{D}$. In addition, we estimate the sharp lower bound of $\mathrm{Re}\, p$ for different choices of h and derive a univalence criteria for functions in $\mathcal{H}$(class of analytic normalized functions) as an application to our results.

1. Introduction

Let $\mathcal{H}[a, n]$ denote the class of analytic functions f defined on the open unit disk $\mathbb{D} = \{z : |z| < 1\}$, of the form $f(z) = a + a_n z^n + a_{n+1}z^{n+1} + \cdots$, where n is a positive integer and $a \in \mathbb{C}$. Let $\mathcal{H} := \{f \in \mathcal{H}[0, 1] : f'(0) = 1\}$ and $\mathcal{S}$ be the subclass of $\mathcal{H}$ consisting of the univalent functions. The Carathéodory class $\mathcal{P}$ consists of the analytic functions p defined on $\mathbb{D}$ of the form $p(z) = 1 + p_1 z + p_2 z^2 + \cdots$, having positive real part in $\mathbb{D}$. Consider two analytic functions f and g in $\mathbb{D}$, we say $f(z)$ is subordinate to $g(z)$, denoted by $f(z) \prec g(z)$, if $f(z) = g(\omega(z))$, where ω is a Schwarz function such that $\omega(0) = 0$ and $|\omega(z)| < 1$. If g is univalent in $\mathbb{D}$, then $f(z) \prec g(z)$ if and only if $f(0) = g(0)$ and $f(\mathbb{D}) \subseteq g(\mathbb{D})$. Ma and Minda [6] unified many subclasses of starlike and convex functions, as follows:

$$\mathcal{S}^*(\phi) = \left\{f \in \mathcal{S} : \frac{zf'(z)}{f(z)} \prec \phi(z)\right\} \text{ and } \mathcal{C}(\phi) = \left\{f \in \mathcal{S} : 1 + \frac{zf''(z)}{f'(z)} \prec \phi(z)\right\}, \tag{1.1}$$

where ϕ is an analytic function having positive real part in $\mathbb{D}$ and $\phi'(0) > 0$ such that $\phi(\mathbb{D})$ is symmetric with respect to the real axis and starlike with respect to $\phi(0) = 1$. For $\phi(z) = (1+z)/(1-z)$, the above classes, respectively reduces to $\mathcal{S}^*$, the class of starlike univalent functions and $\mathcal{C}$, the class of convex univalent functions. Recall that the p-valent function $f(z) = z^p + a_{p+1}z^{p+1} + \cdots$ is convex if and only if

$$\mathrm{Re}\left(1 + \frac{zf''(z)}{f'(z)}\right) > 0, \quad z \in \mathbb{D}. \tag{1.2}$$

This class is extensively studied in [2]. We may now conclude:

Definition 1.1. The function $f \in \mathcal{H}$ is convex univalent in $\mathbb{D}$ if and only if (1.2) holds.

Moreover, the univalence criteria for such functions is illustrated recently in [10] as follows:

Lemma 1.2. [10] *Let the function $q(z) = z + a_2 z^2 + \cdots$ be analytic in $\mathbb{D}$ such that $q(z)q'(z)/z \neq 0$ holds. Then q satisfying $1 + \mathrm{Re}\,(zq''(z)/q'(z)) > \alpha$ in $\mathbb{D}$ is univalent if and only if $-1/2 \leq \alpha \leq 0$.*

2010 *Mathematics Subject Classification.* 30C45, 30C50.

Key words and phrases. Differential subordination, Exact differential equation, Convex functions, best dominant.

Geometrically, it is also known that a function $f(z)$ is convex(starlike) function if it maps $\mathbb{D}$ onto a convex(starlike) domain. A function $f \in \mathcal{H}$ is starlike if and only if it satisfies

$$\operatorname{Re}\left(\frac{zf'(z)}{f(z)}\right) > 0, \quad z \in \mathbb{D}. \tag{1.3}$$

This characterization for starlikeness of functions in $\mathcal{H}$ cannot be extended to $\mathcal{H}[a,1]$. We can see this in case of $f_0(z) = (1+z)/(1-z) \in \mathcal{H}[1,1]$, which maps the unit disk $\mathbb{D}$ onto a convex domain, hence starlike in $\mathbb{D}$ but fails to satisfy (1.3) as $\operatorname{Re}(zf_0'(z)/f_0(z)) = \operatorname{Re}(2z/(1-z^2)) \not> 0$. This reveals that the normalization $f(0) = 0$ is essential, which is not followed by functions in $\mathcal{H}[a,1]$ $(a \neq 0)$. But, the characterization of convex functions $f \in \mathcal{H}$, given in Definition 1.1 can be extended to convex functions in $\mathcal{H}[a,1]$ as follows:

Lemma 1.3. *The function $\tilde{f}(z) = a + a_1 z + a_2 z^2 + \cdots \in \mathcal{H}[a,1]$ is convex univalent in $\mathbb{D}$ if and only if it satisfies*

$$\operatorname{Re}\left(1 + \frac{z\tilde{f}''(z)}{\tilde{f}'(z)}\right) > 0, \quad z \in \mathbb{D}. \tag{1.4}$$

Proof. Consider a function $f(z) = \frac{\tilde{f}(z)-a}{a_1}$, then $f \in \mathcal{H}$ and we get

$$\operatorname{Re}\left(1 + \frac{z\tilde{f}''(z)}{\tilde{f}'(z)}\right) = \operatorname{Re}\left(1 + \frac{zf''(z)}{f'(z)}\right). \tag{1.5}$$

Now, if we assume $\tilde{f}$ is convex univalent in $\mathbb{D}$, then geometrically so is the function $f(z)$. Thus equation (1.4) follows using Definition 1.1 and equation (1.5). Conversely, let equation (1.4) holds, then equation (1.2) holds true using equation (1.5). Thus by Definition 1.1, we get f is convex univalent in $\mathbb{D}$ and geometrically so is $\tilde{f}(z)$. This completes the proof. ∎

The theory of differential subordinations is extensively studied in [8]. For details regarding dominants and best dominant of the differential subordinations, see [8, pp. 16]. For some more work in this direction, one may refer [1, 3, 5]. The set of analytic and univalent functions $\tilde{q}$ on $\overline{\mathbb{D}}\backslash E(\tilde{q})$, where

$$E(\tilde{q}) = \{\zeta \in \partial\mathbb{D} : \lim_{z\to\zeta} \tilde{q}(z) = \infty\},$$

such that $\tilde{q}'(\zeta) \neq 0$ for $\zeta \in \partial\mathbb{D}\backslash E(\widetilde{q})$ is denoted by $\widetilde{Q}$. The theory of admissible functions paved a way in finding the dominants of the solutions of differential subordination altogether with a distinctive approach. The admissible function is defined as follows:

Definition 1.4. [8] Let Ω be a set in $\mathbb{C}$, $\tilde{q} \in \widetilde{Q}$ and n be a positive integer. The class of admissible functions $\Psi_n[\Omega, \tilde{q}]$, consists of the functions $\psi : \mathbb{C}^2 \times \mathbb{D} \to \mathbb{C}$, which satisfy the admissibility condition

$$\psi(r, s; z) \notin \Omega, \tag{1.6}$$

whenever $r = \tilde{q}(\zeta)$, $s = m\zeta\tilde{q}'(\zeta)$, $z \in \mathbb{D}$, $\zeta \in \partial\mathbb{D}\backslash E(\tilde{q})$ and $m \geq n$.

In particular, if Ω is a simply connected domain, $\Omega \neq \mathbb{C}$ and h conformally maps $\mathbb{D}$ onto Ω, the class $\Psi_n[\Omega, \tilde{q}]$ is denoted by $\Psi_n[h, \tilde{q}]$. Recently, in [9], authors have studied the applications of admissibility conditions for various known classes of starlike functions. We consider a special case for the class $\Psi_n[\Omega, \tilde{q}]$, where $\Omega = \{w : \operatorname{Re} w > 0\}$, $\tilde{q}(\mathbb{D}) = \Omega$, $\tilde{q}(0) = 1$, $E(\tilde{q}) = \{1\}$ and $\tilde{q} \in \widetilde{Q}$, then above admissibility condition, given in (1.6) reduces to

$$\psi(\rho i, \sigma; z) \notin \Omega,$$

where ρ and $\sigma \in \mathbb{R}$, $\sigma \leq -(n/2)|1 - i\rho|^2$, $z \in \mathbb{D}$ and $n \geq 1$. In this special case, set the class $\Psi_n[\Omega, q] =: \Psi_n\{1\}$.

Now, we recall that a first order differential equation of the type:

$$M(x,y)dx + N(x,y)dy = 0, \tag{1.7}$$

where $M(x,y)$ and $N(x,y)$ have continuous partial derivatives in some domain $D \subset \mathbb{R}^2$, is called an exact differential equation if and only if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. The solution of such a differential equation (1.7), is given by:

$$\int M(x,y)dx + \int \tilde{N}(y)dy = c, \tag{1.8}$$

where c is the integration constant and $\tilde{N}(y)$ is equal to that part of the expression $N(x,y)$, which is independent of x. This concept we carry to the complex plane in the following definition:

Definition 1.5. Let $p \in \mathcal{H}[a_0, n]$ for some suitable a_0 and ψ be an analytic function satisfying the following differential subordination

$$\psi(p(z), zp'(z)) \prec h(z), \tag{1.9}$$

which further implies $p \prec q$, where p is a solution of (1.9) and q is the best dominant satisfying the differential equation

$$\psi(q(z), nzq'(z)) = h(z). \tag{1.10}$$

Then the differential subordintion (1.9) is said to be an *exact differential subordination* if and only if the following conditions hold:
(i) ψ can be expressed as $\widetilde{M}(z,p) + N(z,p)\ p'(z)$ such that $\frac{\partial \widetilde{M}}{\partial p} = \frac{\partial N}{\partial z}$ for all n.
(ii) For $n = 1$, the equation (1.10) reduces to an exact differential equation:

$$M(z,q)\ dz + N(z,q)\ dq = 0,$$

or equivalently $\frac{\partial M}{\partial q} = \frac{\partial N}{\partial z}$, where $M(z,q) := \widetilde{M}(z,q) - h(z)$.

In the present investigation, we mainly focus on the following exact type of differential subordinations:

$$(\beta p(z) + \gamma)^{-\alpha} \left(\frac{(\beta p(z) + \gamma)}{\beta(1-\alpha)} + zp'(z) \right) \prec h(z), \tag{1.11}$$

and

$$\frac{1}{\sqrt{\gamma\beta}} \arctan\left(\sqrt{\frac{\beta}{\gamma}} p^{1-\alpha}(z) \right) + \left(\frac{1-\alpha}{\beta p^{2(1-\alpha)}(z) + \gamma} \right) \frac{zp'(z)}{p^{\alpha}(z)} \prec h(z), \tag{1.12}$$

where h is a non vanishing convex univalent function in $\mathbb{D}$ with $h(0) = a$ and $p \in \mathcal{H}[a_0, n]$ for appropriate a_0. We denote the expression on left side of (1.11) and (1.12) by $\psi_1(p(z), zp'(z))$ and $\psi_2(p(z), zp'(z))$ respectively. The speciality of the above two differential subordinations is that they generalize a result studied by Hallenbeck and Ruscheweyh [4]. Here we find the dominant and the best dominant q of the solutions of the above differential subordinations, which satisfy

$$\psi_i(q(z), nzq'(z)) = h(z) \quad (i = 1, 2). \tag{1.13}$$

Now we show that the above two differential subordinations are exact. Take $\widetilde{M}(z,p) := \frac{(\beta p+\gamma)^{1-\alpha}}{\beta(1-\alpha)}$ and $N(z,p) := \frac{z}{(\beta p+\gamma)^{\alpha}}$ in (1.11), then we have $\frac{\partial \widetilde{M}}{\partial p} = \frac{\partial N}{\partial z}$. Further, for $n = 1$ the equation (1.13) with $i = 1$, reduces to the following exact differential equation in z and q:

$$M(z,q)\ dz + N(z,q)\ dq = 0 \tag{1.14}$$

as $\frac{\partial M}{\partial q} = \frac{\partial N}{\partial z}$. Thus (1.11) is an exact differential subordination. Since (1.14) is an exact differential equation, we can obtain its solution analogous to (1.8) as follows:

$$\int M(z,q)\ dz + \int \tilde{N}(z)\ dq = c, \tag{1.15}$$

where $\tilde{N}(z)$ is equal to that part of the expression $N(z,q)$, which is independent of q. Upon simplification of the above equation, we get

$$q(z)=\frac{\left(\frac{\beta(1-\alpha)}{nz^{1/n}}\int_0^z h(t)t^{(1/n)-1}dt\right)^{\frac{1}{1-\alpha}}-\gamma}{\beta}. \tag{1.16}$$

Similarly, we can show that $\psi_2(p(z),zp'(z))\prec h(z)$ is an exact differential subordination. Interestingly the q, we obtained in (1.16), coincide with the best dominant obtained by the theory of differential subordinations [8]. This is illustrated in the subsequent section. Further, we show that the dominant obtained in each of the cases is convex univalent by using the concept of admissibility condition (1.6), which ultimately generalizes and improves some of the earlier known results. We also give some specific examples to justify our claims. In addition, we estimate the best dominant for different choices of h. Also, we find a criteria for univalence of functions in $\mathcal{H}$ as an application to our results. Further, using this univalence criteria, we obtain a few more results.

In what follows, let us presume convex to mean convex univalent.

2. Main Results

We need the following Hallenbeck and Ruscheweyh [4] result with $\gamma=1$ to prove our main results.

Lemma 2.1. [4] *Let h be convex in $\mathbb{D}$ having $h(0)=a$. If $P\in\mathcal{H}[a,n]$ satisfies*

$$P(z)+zP'(z)\prec h(z),$$

then $P(z)\prec Q(z)\prec h(z)$, where

$$Q(z)=\frac{1}{nz^{1/n}}\int_0^z h(t)t^{(1/n)-1}dt, \tag{2.1}$$

is convex and is the best $(a,n)-$dominant.

The following theorem deals in finding the best dominant of the exact type differential subordination:

Theorem 2.2. *Let $\beta(\neq 0)$, γ be complex numbers, $-1\leq\alpha\leq 0$ and h be a non-vanishing convex function with $h(0)=a$. Let $a_0=((\beta(1-\alpha)a)^{1/(1-\alpha)}-\gamma)/\beta$ and $p\in\mathcal{H}[a_0,n]$ satisfies*

$$\frac{(\beta p(z)+\gamma)^{1-\alpha}}{\beta(1-\alpha)}+\frac{zp'(z)}{(\beta p(z)+\gamma)^{\alpha}}\prec h(z), \tag{2.2}$$

then $p(z)\prec q(z)\prec H(z)$, where $H(z):=((\beta(1-\alpha)h(z))^{\frac{1}{1-\alpha}}-\gamma)/\beta$,

$$q(z)=\frac{\left(\frac{\beta(1-\alpha)}{nz^{1/n}}\int_0^z h(t)t^{(1/n)-1}dt\right)^{\frac{1}{1-\alpha}}-\gamma}{\beta}, \tag{2.3}$$

are convex in $\mathbb{D}$ and q is the best (a_0,n)- dominant.

Proof. We first prove that q and H are convex in $\mathbb{D}$. Since the function h here satisfies the conditions of h of Lemma 2.1, the expression (2.3) becomes

$$q(z)=\frac{(\beta(1-\alpha)Q(z))^{1/(1-\alpha)}-\gamma}{\beta}, \tag{2.4}$$

where Q is given by (2.1). Thus q is well-defined and analytic in $\mathbb{D}$ as $\beta\neq 0$. Let $P(z)=\frac{(\beta p(z)+\gamma)^{1-\alpha}}{\beta(1-\alpha)}$, then $P\in\mathcal{H}[a,n]$ and the subordination (2.2) becomes $P(z)+zP'(z)\prec h(z)$. Now

from Lemma 2.1, we have $P(z) \prec Q(z) \prec h(z)$, which implies $(\beta p(z)+\gamma)^{1-\alpha} \prec \beta(1-\alpha)Q(z) \prec \beta(1-\alpha)h(z)$. Since $h(z)$ does not vanish in $\mathbb{D}$, we have $q'(z) \neq 0$ and therefore $1+zq''(z)/q'(z) =: s(z)$ is analytic in $\mathbb{D}$. Logarithmic differentiation of (2.4) yields

$$1+\frac{zQ''(z)}{Q'(z)} = 1+\frac{zq''(z)}{q'(z)} + A(z) =: \tilde{\psi}(s(z); z), \tag{2.5}$$

where

$$A(z) = -\alpha\beta\frac{zq'(z)}{\beta q(z)+\gamma} = \frac{\alpha}{\alpha-1}\frac{zQ'(z)}{Q(z)}$$

and we can write (2.5) as $\tilde{\psi}(r) = r + A(z)$. By Lemma 2.1, the function Q is convex, thus from Lemma 1.3, we have $\operatorname{Re}\tilde{\psi}(s(z)) > 0$. Now to prove q is convex, by Lemma 1.3, it suffices to show $\operatorname{Re} s(z) > 0$. To accomplish this task, we use [8, Theorem 2.3i, p. 35] and so we need to show that $\tilde{\psi} \in \Psi_n\{1\}$. Further, by using the admissibility condition for the same, it is equivalent to prove that $\operatorname{Re}(\tilde{\psi}(\rho i; z)) \not> 0$. From equation (2.5), we obtain

$$\operatorname{Re}(\tilde{\psi}(\rho i)) = \operatorname{Re}(\rho i + A(z)) = \operatorname{Re} A(z).$$

Using the fact that $Q(z) \neq 0$ for all $z \in \mathbb{D}$, we get $zQ'(z)/Q(z)$ is analytic in $\mathbb{D}$ and since $P(z) \prec Q(z)$, we have $Q(0) = a(\neq 0)$, which implies

$$\left.\frac{zQ'(z)}{Q(z)}\right|_{z=0} = 0 \tag{2.6}$$

and also we have $Q'(z) \neq 0$ for all z, therefore the quantity, given in (2.6) lies on either side of the imaginary axis. So clearly we have $\operatorname{Re} A(z) \not> 0$. Thus, q is convex in $\mathbb{D}$. Similarly, we can prove the function H is convex in $\mathbb{D}$, since $zh'(z)/h(z)$ $(h(0) = a)$ and $zQ'(z)/Q(z)$ behave alike. We now proceed to show $p(z) \prec q(z) \prec H(z)$ and q is the best (a_0, n)- dominant. Let $\psi(r,s) := \frac{(\beta r+\gamma)^{1-\alpha}}{\beta(1-\alpha)} + \frac{s}{(\beta r+\gamma)^{\alpha}}$, which corresponds to left hand side of the subordination (2.2). First we show $p \prec H$ by proving $\psi \in \Psi_n[h, H]$, or equivalently

$$\psi_0 := \psi(H(\zeta), m\zeta H'(\zeta)) = \frac{(\beta H(\zeta)+\gamma)^{1-\alpha}}{\beta(1-\alpha)} + \frac{m\zeta H'(\zeta)}{(\beta H(\zeta)+\gamma)^{\alpha}} \notin h(\mathbb{D}),$$

whenever $|\zeta| = 1$ and $m \geq n$. From hypothesis, replacing H by its expression in terms of h, in ψ_0, we obtain

$$\left|\arg\left(\frac{\psi_0 - h(\zeta)}{\zeta h'(\zeta)}\right)\right| = |\arg m| < \pi/2.$$

As the function $h(\mathbb{D})$ is convex and $m \geq 1$, $h(\zeta) \in h(\partial\mathbb{D})$ and $\zeta h'(\zeta)$ is the outer normal to $h(\partial\mathbb{D})$ at $h(\zeta)$, we arrive at the conclusion that $\psi_0 \notin h(\mathbb{D})$ and therefore we obtain $p(z) \prec H(z)$. Now, we show $p \prec q$. A simple calculation shows that q, given by (2.3) satisfies the following differential equation

$$\frac{(\beta q(z)+\gamma)^{1-\alpha}}{\beta(1-\alpha)} + \frac{nzq'(z)}{(\beta q(z)+\gamma)^{\alpha}} = \psi(q(z), nzq'(z)) = h(z). \tag{2.7}$$

Now we apply [8, Theorem 2.3f, pp.32] to show q is the best dominant. Without loss of generality, we can assume h and q are analytic and univalent on $\overline{\mathbb{D}}$ and $q'(\zeta) \neq 0$ for $|\zeta| = 1$, which shows that $q \in \widetilde{Q}$. To complete the proof, we now show $\psi \in \Psi_n[h, q]$. This is equivalent to show that

$$\psi_{00} := \psi(q(\zeta), m\zeta q'(\zeta)) = \frac{(\beta q(\zeta)+\gamma)^{1-\alpha}}{\beta(1-\alpha)} + \frac{m\zeta q'(\zeta)}{(\beta q(\zeta)+\gamma)^{\alpha}} \notin h(\mathbb{D}),$$

whenever $|\zeta| = 1$ and $m \geq n$. Using equations (2.3) and (2.7), we obtain

$$Q(\zeta) + \frac{m}{n}\left(h(\zeta) - Q(\zeta)\right) = \psi_{00}.$$

Since we have $Q(z) \prec h(z)$ from Lemma 2.1 and also we know $m/n \geq 1$, it is easy to observe that $\psi_{00} \notin h(\mathbb{D})$. This completes the proof. ∎

Let $n = 1$, then the best dominant q, given in (2.3) is same as the solution (1.16) obtained by solving its associated exact differential equation. By taking $\beta = 1$ and $\gamma = 0$ in Theorem 2.2, we get the following result:

Corollary 2.3. *Let $-1 \leq \alpha \leq 0$, h be non vanishing convex function in $\mathbb{D}$ with $h(0) = a$ and $a_0 = ((1-\alpha)a)^{1/(1-\alpha)}$. If $p \in \mathcal{H}[a_0, n]$ satisfies*

$$\frac{p^{1-\alpha}(z)}{1-\alpha} + \frac{zp'(z)}{p^{\alpha}(z)} \prec h(z), \tag{2.8}$$

then $p(z) \prec q(z) \prec H(z)$, where $H(z) := ((1-\alpha)h(z))^{\frac{1}{1-\alpha}}$,

$$q(z) = \left(\frac{(1-\alpha)}{nz^{1/n}} \int_0^z h(t)t^{(1/n)-1}dt\right)^{\frac{1}{1-\alpha}}. \tag{2.9}$$

are convex in $\mathbb{D}$ and q is the best (a_0, n)- dominant.

In Corollary 2.3, if we choose $\alpha = -1$ and h as any convex function with $h(0) = 1$ and $\operatorname{Re} h(z) > 0$, then we obtain the following result:

Corollary 2.4. *Let h be a convex function with $h(0) = 1$ and $\operatorname{Re} h(z) > 0$. If $p \in \mathcal{H}[1, n]$ satisfies the following*

$$p^2(z) + 2zp(z)p'(z) \prec h(z),$$

then

$$p(z) \prec q(z) = \sqrt{Q(z)}, \text{ where } Q(z) = \frac{1}{nz^{1/n}} \int_0^z h(t)t^{(1/n)-1}dt. \tag{2.10}$$

The function q is convex and best $(1, n)$- dominant.

Remark 2.5. The Corollary 2.4 improves the already known result of Miller and Mocanu [8, Theorem 3.1e., p. 77], by additionally establishing q, given in (2.10) is convex in $\mathbb{D}$ apart from being the best dominant. Moreover the Corollary 2.3 generalizes the known result [8, Theorem 3.1e., p. 77] for all α $(-1 \leq \alpha \leq 0)$ and extends the same by proving the convexity of the function q as well.

For requisite knowledge about the Hypergeometric functions, one may refer [8, p. 5-7]. We now obtain the following corollary for different choices of h in Theorem 2.2, when β is a real number

Corollary 2.6. *Let $-1 \leq \alpha \leq 0$, $\beta > 0$ and γ be a complex number. If $p \in \mathcal{H}[(\beta(1-\alpha))^{1/(1-\alpha)} - \gamma)/\beta, n]$ and $\psi_1(p(z), zp'(z))$, given in (1.11) satisfies any of the following:*

(i) *$\psi_1 \prec (1+Az)/(1+Bz)$, $-1 \leq B < A \leq 1$, with*
$\lambda_1 := {}_2F_1\left(1, 1/n, 1+1/n; B\right) - \frac{A}{n+1}\, {}_2F_1\left(1, 1+1/n, 2+1/n; B\right)$

(ii) *$\psi_1 \prec e^{\mu z}$, $(|\mu| \leq 1)$, with $\lambda_2 := {}_1F_1\left(1/n, 1/n+1; -\mu\right)$*

(iii) *$\psi_1 \prec \sqrt{1+\kappa z}$, $\kappa \in [0, 1]$, with $\lambda_3 := {}_2F_1\left(-1/2, 1/n, 1+1/n; \kappa\right)$*

(iv) *$-\frac{\rho_2\pi}{2} < \arg \psi_1 < \frac{\rho_1\pi}{2}$, where $\rho = \frac{\rho_1-\rho_2}{\rho_1+\rho_2}$, $\rho' = \frac{\rho_1+\rho_2}{2}$, $0 < \rho_1, \rho_2 \leq 1$, and $c = e^{\rho\pi i}$, with*
$\lambda_4 := \displaystyle\sum_{k=0}^{\infty} \left(\frac{\binom{\rho'}{k}(-c)^k {}_2F_1(\rho', 1/n+k, 1+1/n+k; -1)}{1+nk} \right),$

then respectively for the above parts $(i)-(iv)$*, we have* $\operatorname{Re} p(z) > \zeta_i(\alpha,\beta,\gamma,\lambda_i)$*, where*

$$\zeta_i(\alpha,\beta,\gamma,\lambda_i) = \operatorname{Re}(((\beta(1-\alpha)\lambda_i)^{1/(1-\alpha)} - \gamma)/\beta)$$

with i *assuming appropriate integral value between* 1 *and* 4*. The result is sharp.*

Proof. Let $h(z) = (1+Az)/(1+Bz)$, $e^{\mu z}$, $\sqrt{1+\kappa z}$ and $((1+cz)/(1-z))^{\rho'}$, respectively for the parts $(i)-(iv)$. Then from Theorem 2.2, we have $\operatorname{Re} p(z) > \min_{|z|\le 1} \operatorname{Re} q(z)$, where q is given by (2.4), whenever either of the above parts $(i)-(iv)$ holds. Therefore, it suffices to find the minimum value of real part of Q in each of the parts, where Q is given by (2.1).
(i) We have $h(z) = (1+Az)/(1+Bz)$, then from equation (2.1) we get

$$\begin{aligned} Q(z) &= \frac{1}{nz^{1/n}} \int_0^z \left(\frac{1+At}{1+Bt}\right) t^{(1/n)-1} dt \\ &= \frac{1}{n} \int_0^1 (1+Btz)^{-1} t^{(1/n)-1} dt + \frac{Az}{n} \int_0^1 (1+Btz)^{-1} t^{1/n} dt \\ &= {}_2F_1\left(1, 1/n, 1+1/n; -Bz\right) + \frac{Az}{n+1}\, {}_2F_1\left(1, 1+1/n, 2+1/n; -Bz\right). \end{aligned}$$

Now, we have $\min_{|z|\le 1} \operatorname{Re} Q(z) = Q(-1) = \lambda_1$, where

$$\lambda_1 := {}_2F_1\left(1, 1/n, 1+1/n; B\right) - \frac{A}{n+1}\, {}_2F_1\left(1, 1+1/n, 2+1/n; B\right).$$

Hence the result.
(ii) We have $h(z) = \exp(\mu z)$, then from equation (2.1) we obtain

$$\begin{aligned} Q(z) &= \frac{1}{nz^{1/n}} \int_0^z e^{\mu t} t^{1/n-1} dt \\ &= \tfrac{1}{n} \int_0^1 e^{\mu t z} t^{1/n-1} dt \\ &= {}_1F_1\left(1/n, 1/n+1; \mu z\right). \end{aligned}$$

Clearly, $\min_{|z|\le 1} \operatorname{Re} Q(z) = Q(-1) = \lambda_2$, where

$$\lambda_2 := {}_1F_1\left(1/n, 1/n+1; -\mu\right).$$

This completes the proof for part (ii).
(iii) Let $h(z) = \sqrt{1+\kappa z}$, then from equation (2.1), we have

$$\begin{aligned} Q(z) &= \frac{1}{nz^{1/n}} \int_0^z \sqrt{1+\kappa t}\; t^{(1/n)-1} dt \\ &= \frac{1}{n} \int_0^1 \sqrt{1+\kappa t z}\; t^{(1/n)-1} dt \\ &= {}_2F_1\left(-1/2, 1/n, 1+1/n; -\kappa z\right). \end{aligned}$$

Thus, we have $\min_{|z|\le 1} \operatorname{Re} Q(z) = Q(-1) = \lambda_3$, where

$$\lambda_3 := {}_2F_1\left(-1/2, 1/n, 1+1/n; \kappa\right)$$

and that completes the proof for part (iii).

(iv) We have $h(z) = \left(\frac{1+cz}{1-z}\right)^{\rho'}$, then from equation (2.1), we have

$$\begin{aligned} Q(z) &= \frac{1}{nz^{1/n}} \int_0^z \left(\frac{1+ct}{1-t}\right)^{\rho'} t^{(1/n)-1} dt \\ &= \frac{1}{nz^{1/n}} \int_0^z \left(\sum_{k=0}^{\infty} \binom{\rho'}{k} c^k t^{(1/n+k-1)} (1-t)^{-\rho'} \right) dt \\ &= \frac{1}{nz^{1/n}} \sum_{k=0}^{\infty} \left(\binom{\rho'}{k} c^k \int_0^1 (tz)^{(1/n+k-1)} (1-tz)^{-\rho'} z dt \right) \\ &= \frac{1}{n} \sum_{k=0}^{\infty} \left(\binom{\rho'}{k} (cz)^k \int_0^1 t^{(1/n+k-1)} (1-tz)^{-\rho'} dt \right) \\ &= \sum_{k=0}^{\infty} \left(\frac{\binom{\rho'}{k} (cz)^k {}_2F_1(\rho', 1/n+k, 1+1/n+k; z)}{1+nk} \right). \end{aligned}$$

Now, $\min_{|z|\leq 1} \operatorname{Re} Q(z) = Q(-1) = \lambda_4$, where

$$\lambda_4 := \sum_{k=0}^{\infty} \left(\frac{\binom{\rho'}{k} (-c)^k {}_2F_1(\rho', 1/n+k, 1+1/n+k; -1)}{1+nk} \right).$$

Hence the result. ∎

Remark 2.7. If A and B are replaced with each other in Corollary 2.6(i), then $\min_{|z|\leq 1} \operatorname{Re} Q(z) = Q(1)$. Further by taking $A = 2a - 1$ $(a \in [0, 1))$, $B = 1$, $\alpha = -1$, $\beta = 1$ and $\gamma = 0$, the above result reduces to the result of Miller and Mocanu [8, Corollary 3.1e.1, p. 79].

In the following theorem, we deal with another exact type differential subordination.

Theorem 2.8. *Let β, $\gamma(\neq 0)$ be complex numbers, $-1 \leq \alpha \leq 0$, h be a non vanishing convex function in $\mathbb{D}$ with $h(0) = a$, $|\sqrt{\gamma\beta} h(z)| < \pi/2$ and $a_0 = (\sqrt{\gamma/\beta} \tan(\sqrt{\gamma\beta} a))^{1/(1-\alpha)}$. If $p \in \mathcal{H}[a_0, n]$ satisfies*

$$\frac{1}{\sqrt{\gamma\beta}} \arctan\left(\sqrt{\frac{\beta}{\gamma}} p^{1-\alpha}(z)\right) + \left(\frac{1-\alpha}{\beta p^{2(1-\alpha)}(z) + \gamma}\right) \frac{zp'(z)}{p^{\alpha}(z)} \prec h(z), \tag{2.11}$$

then $p(z) \prec q(z) \prec H(z)$, where $H(z) = (\sqrt{\gamma/\beta} \tan(\sqrt{\gamma\beta} h(z)))^{1/(1-\alpha)}$ and

$$q(z) = \left(\sqrt{\frac{\gamma}{\beta}} \tan\left(\frac{\sqrt{\gamma\beta}}{nz^{1/n}} \int_0^z h(t) t^{(1/n)-1} dt\right)\right)^{\frac{1}{1-\alpha}} \tag{2.12}$$

are convex in $\mathbb{D}$ and q is the best $(a_0, n)-$dominant.

Proof. We first show that q and H are convex in $\mathbb{D}$. As the function h here satisfies the condition of h of Lemma 2.1, thus q reduces to

$$q(z) = \left(\sqrt{\frac{\gamma}{\beta}} \tan\left(\sqrt{\gamma\beta} Q(z)\right)\right)^{\frac{1}{1-\alpha}}, \tag{2.13}$$

where Q is given by (2.1) and the function q is well defined and analytic in $\mathbb{D}$. Let $P(z) = (1/\sqrt{\gamma\beta}) \arctan\left(\sqrt{\frac{\beta}{\gamma}} p^{1-\alpha}(z)\right)$, then the subordination (2.11) reduces to $P(z) + zP'(z) \prec h(z)$ and clearly $P \in \mathcal{H}[a, n]$. Now from Lemma 2.1, we have $P(z) \prec Q(z) \prec h(z)$, which is equivalent to

$\arctan\left(\sqrt{\frac{\beta}{\gamma}}p^{1-\alpha}(z)\right) \prec \sqrt{\gamma\beta}Q(z) \prec \sqrt{\gamma\beta}h(z)$. We have $q'(z) \neq 0$ as $h(z) \neq 0$ for all $z \in \mathbb{D}$ and therefore $1 + zq''(z)/q'(z) =: g(z)$ is analytic in $\mathbb{D}$. The logarithmic differentiation of the function q, given in (2.13) yields

$$1 + \frac{zQ''(z)}{Q'(z)} = 1 + \frac{zq''(z)}{q'(z)} + B(z) =: \tilde{\psi}(g(z); z), \tag{2.14}$$

where

$$\begin{aligned} B(z) &= -\left(\alpha + 2(1-\alpha)\beta\frac{q^{2(1-\alpha)}(z)}{\gamma + \beta q^{2(1-\alpha)}(z)}\right)\frac{zq'(z)}{q(z)} \\ &= \frac{\alpha}{\alpha - 1}\left(\sum_{k=0}^{\infty}(-1)^k\frac{(2\sqrt{\gamma\beta}Q(z))^{2k}}{(2k+1)!}\right)^{-1}\left(\frac{zQ'(z)}{Q(z)}\right) \\ &\quad - 2\sqrt{\gamma\beta}\tan(\sqrt{\gamma\beta}Q(z))zQ'(z). \end{aligned} \tag{2.15}$$

and we can write $\tilde{\psi}(r) = r + B(z)$. By Lemma 2.1, the function Q is convex, thus from Lemma 1.3, we have $\operatorname{Re}\tilde{\psi}(g(z)) > 0$. In order to prove q is convex, by Lemma 1.3, it suffices to show $\operatorname{Re} g(z) > 0$. Now, to achieve this, we use [8, Theorem 2.3i, p. 35], so then we only need to show $\tilde{\psi} \in \Psi_n\{1\}$. Further, by using the admissibility condition for the same, it is equivalent to show that $\operatorname{Re}(\tilde{\psi}(\rho i; z)) \not\geq 0$. From equation (2.14), we obtain

$$\operatorname{Re}(\psi(\rho i)) = \operatorname{Re}(\rho i + B(z)) = \operatorname{Re} B(z).$$

Now, using the fact $Q(z) \neq 0$ for all $z \in \mathbb{D}$, we get B is analytic in $\mathbb{D}$. Since $P(z) \prec Q(z)$, we have $Q(z) \in \mathcal{H}[a, n]$, thus

$$\left(\sum_{k=0}^{\infty}(-1)^k\frac{(2\sqrt{\gamma\beta}Q(z))^{2k}}{(2k+1)!}\right)^{-1}\left(\frac{zQ'(z)}{Q(z)}\right)\Big|_{z=0} = 0 \tag{2.16}$$

and

$$\tan(\sqrt{\gamma\beta}Q(z))zQ'(z)|_{z=0} = 0, \tag{2.17}$$

which ensures $B(z) \equiv 0$ or lies on either side of the imaginary axis. Therefore, from equation (2.15), we have $\operatorname{Re} B(z) \not\geq 0$, which proves q is convex in $\mathbb{D}$. Now similarly, we can prove the function H is convex in $\mathbb{D}$. We now show $p(z) \prec q(z) \prec H(z)$ and q is the best (a_0, n)- dominant. Now, for this we assume

$$\psi(r, s) := \frac{1}{\sqrt{\gamma\beta}}\arctan\left(\sqrt{\frac{\beta}{\gamma}}r^{1-\alpha}\right) + \frac{(1-\alpha)s}{r^{\alpha}(\beta r^{2(1-\alpha)} + \gamma)},$$

which corresponds to left hand side of the subordination (2.11). Now first we show $p \prec H$ by proving $\psi \in \Psi_n[h, H]$, or equivalently

$$\begin{aligned} &\frac{1}{\sqrt{\gamma\beta}}\arctan\left(\sqrt{\frac{\beta}{\gamma}}H^{1-\alpha}(\zeta)\right) + \frac{(1-\alpha)m\zeta H'(\zeta)}{H^{\alpha}(\zeta)(\beta H^{2(1-\alpha)}(\zeta) + \gamma)} \\ &\quad = \psi(H(\zeta), m\zeta H'(\zeta)) =: \psi_0 \notin h(\mathbb{D}), \end{aligned}$$

whenever $|\zeta| = 1$ and $m \geq n$. Now, replacing H by its expression in terms of h in ψ_0, we get

$$\left|\arctan\left(\frac{\psi_0 - h(\zeta)}{\zeta h'(\zeta)}\right)\right| = |\arg m| < \frac{\pi}{2}.$$

Since the function h is convex, $h(\zeta) \in h(\partial\mathbb{D})$ and $\zeta h'(\zeta)$ is the outer normal to $h(\partial\mathbb{D})$ at $h(\zeta)$, thus we obtain $\psi_0 \notin h(\mathbb{D})$, which further implies $p(z) \prec H(z)$. A computation shows that q, given by (2.12) satisfies the differential equation

$$\begin{aligned} h(z) &= \frac{1}{\sqrt{\gamma\beta}} \arctan\left(\sqrt{\frac{\beta}{\gamma}} q^{1-\alpha}(z)\right) + \frac{(1-\alpha)nzq'(z)}{q^{\alpha}(z)(\beta q^{2(1-\alpha)}(z)+\gamma)} \\ &=: \psi(q(z), nzq'(z)). \end{aligned} \tag{2.18}$$

Now we apply [8, Theorem 2.3f, p. 32] to show q is the best dominant. Without loss of generality, we can assume h and q are analytic and univalent on $\overline{\mathbb{D}}$ and $q'(\zeta) \neq 0$ for $|\zeta| = 1$, which shows that $q \in \widetilde{Q}$. In order to complete the proof, it suffices to show $\psi \in \Psi_n[h, q]$, which is equivalent to show that

$$\begin{aligned} &\frac{1}{\sqrt{\gamma\beta}} \arctan\left(\sqrt{\frac{\beta}{\gamma}} q^{1-\alpha}(\zeta)\right) + \frac{(1-\alpha)m\zeta q'(\zeta)}{q^{\alpha}(\zeta)(\beta q^{2(1-\alpha)}(\zeta)+\gamma)} \\ &\quad = \psi(q(\zeta), m\zeta q'(\zeta)) =: \psi_{00} \notin h(\mathbb{D}), \end{aligned}$$

whenever $|\zeta| = 1$ and $m \geq n$. The equations (2.1) and (2.18) yield

$$Q(\zeta) + \frac{m}{n}(h(\zeta) - Q(\zeta)) = \psi_{00}.$$

As we have $m/n \geq 1$ and from Lemma 2.1, we have $Q(z) \prec h(z)$, therefore evidently, we get $\psi_{00} \notin h(\mathbb{D})$. This completes the proof. ∎

Taking $n = 1$ in Theorem 2.8, equation (2.18) reduces to the differential equation (1.14) with $M(z,q) := (1/\sqrt{\gamma\beta}) \arctan\left(\sqrt{\frac{\beta}{\gamma}} q^{1-\alpha}\right) - h(z)$ and $N(z,q) := \frac{(1-\alpha)z}{q^{\alpha}(\beta q^{2(1-\alpha)}+\gamma)}$. Further, we have $\frac{\partial M}{\partial q} = \frac{\partial N}{\partial z}$, which shows that the differential equation, given in (2.18) is exact. Consequently, the equation (1.8) yields the solution for equation (2.18) whenever $n = 1$ as follows:

$$(1/\sqrt{\gamma\beta}) z \arctan\left(\sqrt{\frac{\beta}{\gamma}} q^{1-\alpha}(z)\right) - \int_0^z h(t)dt = 0,$$

which coincides with (2.12) with $n = 1$. We now obtain the following corollary for different choices of h in Theorem 2.8, when β is a real number.

Corollary 2.9. *Let $-1 \leq \alpha \leq 0$, $\beta > 0$ and $\gamma(\neq 0)$ be a complex number. If $p \in \mathcal{H}[a_0, n]$, where $a_0 = (\sqrt{\gamma/\beta}\tan(\sqrt{\gamma\beta}a))^{1/(1-\alpha)}$ and $\psi_2(p(z), zp'(z))$, given by (1.12) satisfies any of the following:*

(i) *$\psi_2 \prec (1+Az)/(1+Bz)$, where $-1 \leq B < A \leq 1$ such that $|\sqrt{\gamma\beta}(1+Az)/(1+Bz)| < \pi/2$*

(ii) *$\psi_2 \prec e^{\mu z}$, where $|\mu| \leq 1$ such that $|\sqrt{\gamma\beta}e^{\mu z}| < \pi/2$*

(iii) *$\psi_2 \prec \sqrt{1+\kappa z}$, where $\kappa \in [0,1]$ such that $|\sqrt{\gamma\beta}\sqrt{1+\kappa z}| \leq \pi/2$*

(iv) *$-\rho_2\pi/2 < \arg\psi_2 < \rho_1\pi/2$, $\rho = \frac{\rho_1-\rho_2}{\rho_1+\rho_2}$, where $0 < \rho_1, \rho_2 \leq 1$, and $c = e^{\rho\pi i}$ such that $|\sqrt{\gamma\beta}((1+cz)/(1-z))^{\rho'}| < \pi/2$, $(\rho' = (\rho_1+\rho_2)/2)$,*

then respectively for the above parts $(i)-(iv)$, we have $\operatorname{Re} p(z) > \xi_i(\alpha,\beta,\gamma,\lambda_i)(i = 1,2,3,4)$, where

$$\xi_i(\alpha,\beta,\gamma,\lambda_i) = \operatorname{Re}\left(\sqrt{\gamma/\beta}\tan\left(\sqrt{\gamma\beta}\lambda_i\right)\right)^{\frac{1}{1-\alpha}}$$

and λ_i is same as in the Corollary 2.6. The result is sharp.

Proof. The proof of each part of this Corollary is on the similar lines of the proof of the Corollary 2.6. Therefore, omitted. ∎

3. Specific Examples:

We provide here below some illustrations to our results.

Example 1. If $p \in \mathcal{H}[1,1]$ satisfies

$$p(z) + zp'(z) \prec \frac{1+z}{1-z}, \tag{3.1}$$

then $p(z) \prec q(z) := -2\frac{\log(1-z)}{z} - 1 \prec \frac{1+z}{1-z}$. Moreover, q is the best $(1,1)$- dominant and is convex in $\mathbb{D}$.

Proof. By taking $h(z) = \frac{1+z}{1-z}$, $n = 1$ and $\alpha = 0$ in Corollary 2.3, we obtain $Q(z) = q(z) = -2(\log(1-z))/z - 1$. Now the assertion follows at once from Corollary 2.3 as $h(0) = 1$ and $h(z) \neq 0$ for all $z \in \mathbb{D}$. Moreover, the quantities $zQ'(z)/Q(z)$ and $zH'(z)/H(z)$ lie on either side of the imaginary axis as proved in the proof of Theorem 2.2. ∎

Remark 3.1. The above result can also be stated as if $p \in \mathcal{H}[1,1]$ satisfies

$$\operatorname{Re}(p(z) + zp'(z)) > 0, \tag{3.2}$$

then

$$\begin{aligned}\operatorname{Re} p(z) &> 2\log 2 - 1 \\ &= {}_2F_1\left(1,1,2;-1\right) - \frac{1}{2}\,{}_2F_1\left(1,2,3;-1\right),\end{aligned}$$

which is also a special case of Corollary 2.6(i) with $A = 1$, $B = -1$, $\beta = 1$, $\gamma = 0$ and $\alpha = 0$.

Example 2. Let $a_0 = 2((2/3)^{3/4} - 1)$. If $p \in \mathcal{H}[a_0, 1]$ satisfies the differential subordination

$$\frac{3(p(z)/2+1)^{4/3}}{2} + zp'(z)(p(z)/2+1)^{1/3} \prec \exp(z), \tag{3.3}$$

then $p(z) \prec q(z) \prec H(z)$, where $H(z) = 2((2e^z/3)^{3/4} - 1)$ and $q(z) = 2((2(e^z-1)/(3z))^{3/4} - 1)$. Moreover, the function q is convex and is the best $(a_0, 1)$- dominant.

Proof. Let $h(z) = \exp(z)$, $n = 1$, $\alpha = -1/3$, $\beta = 1/2$ and $\gamma = 1$ in Theorem 2.2, then we obtain $Q(z) = (e^z - 1)/z$. Now h here with the given constants satisfy the hypothesis of Theorem 2.2, the assertion follows from the same. Moreover, in accordance with the proof of Theorem 2.2, we observe that

$$\operatorname{Re}\frac{zQ'(z)}{Q(z)} = \operatorname{Re}\left(\frac{1 + e^z(z-1)}{e^z - 1}\right) \not> 0$$

and similarly the quantity $\operatorname{Re} zH'(z)/H(z) = \operatorname{Re} z \not> 0$, therefore both the quantities lie on either side of the imaginary axis. ∎

Example 3. Let $a_0 = ((0.5)\tan(0.5))^{3/5}$. If $p \in \mathcal{H}[a_0, 1]$ satisfies

$$2\arctan\left(2p^{5/3}(z)\right) + \frac{20}{3}\left(\frac{zp^{2/3}(z)p'(z)}{1 + 4p^{10/3}(z)}\right) \prec \frac{2+z}{2-z}, \tag{3.4}$$

then

$$p(z) \prec q(z) := \left(\tan\left(\frac{-1}{2} - \frac{2}{z}\log\left(\frac{2-z}{2}\right)\right)\Big/2\right)^{3/5}. \tag{3.5}$$

Moreover, the function q is the best $(a_0, 1)$-dominant and is convex in $\mathbb{D}$.

Proof. Let $h(z) = \frac{2+z}{2-z}$, $n = 1$, $\alpha = -2/3$, $\beta = 1$ and $\gamma = 1/4$ in Theorem 2.8, then we obtain $Q(z) = -1 - ((4/z)\log((2-z)/z))$ and $q(z)$ is given by (3.5). As the function h and all the constants satisfy the hypothesis of Theorem 2.8, the result follows at once. Also, we observe that the quantity in (2.15) lie on either side of the imaginary axis, which is in accordance with the proof of Theorem 2.8 ∎

3.1. **Applications to Univalent function.** We obtain a sufficient condition for univalence of f in the following theorem.

Theorem 3.2. *Let $f \in \mathcal{H}$. If f satisfies*

$$\operatorname{Re}(f'(z) + zf''(z)) > 0, \tag{3.6}$$

then f is univalent in $\mathbb{D}$ and in fact, $\operatorname{Re} f' > 2\log 2 - 1$.

Proof. Let $p(z) = f'(z)$, then we have $p \in \mathcal{H}[1,1]$. We observe that equation (3.6) becomes equivalent to equation (3.1) for $p(z) = f'(z)$. Since p satisfies the hypothesis of Example 1, we obtain $\operatorname{Re} f'(z) > 2\log 2 - 1 > 0$. Thus, Noshiro-Warschawski result yields the function f is univalent. ∎

The result by Miller Mocanu [7] with $n = 1$ is as follows:

Corollary 3.3. [7] *Let $a \in [0,1)$ and $\eta = \eta(a)$ be defined as*

$$\eta = \left(2(1-a)\beta(1) + (2a-1)\right)^{1/2},$$

where $\beta(x) = \int_0^1 \frac{t^{x-1}}{1+t}dt$. If $f \in \mathcal{H}$, then

$$\operatorname{Re} f'(z) > a \Rightarrow \operatorname{Re}\sqrt{\frac{f(z)}{z}} > \eta(a). \tag{3.7}$$

Corollary 3.4. *Let $f \in \mathcal{H}$ and $a = 2\log 2 - 1$. If f satisfies the equation (3.6), then*
(i) $\operatorname{Re}\sqrt{\frac{f(z)}{z}} > \eta(a)$.
(ii) $\operatorname{Re}\frac{f(z)}{z} > a$.

Proof. (i) The proof follows directly from Theorem 3.2 and Corollary 3.3.
(ii) Let $p(z) = f(z)/z$, then we have $p \in \mathcal{H}[1,1]$. Thus Remark 3.1 yields $\operatorname{Re} f'(z) > 0 \Rightarrow \operatorname{Re}(f(z)/z) > a$. Thus, the result follows now using Theorem 3.2. ∎

Acknowledgment. The work presented here is supported by a Research Fellowship from the Department of Science and Technology, New Delhi.

References

[1] R. M. Ali, N. E. Cho, V. Ravichandran and S. S. Kumar, Differential subordination for functions associated with the lemniscate of bernoulli, Taiwan. J. Math. **16** (2012), no. 3, 1017–1026.
[2] R. M. Ali, V. Ravichandran and S. K. Lee, Subclasses of multivalent starlike and convex functions, Bull. Belgian Math. Soc.-Simon Stevin. **16** (2009), no. 3, 385–394.
[3] S. Banga and S. S. Kumar, Applications of Applications of differential subordinations to certain classes of starlike functions, J. Korean Math. Soc. **57** (2020), no. 2, 331–357.
[4] D. J. Hallenbeck and S. Ruscheweyh, Subordination by convex functions, Proc. Amer. Math. Soc. **52** (1975), 191–195.
[5] S. Kumar and V. Ravichandran, Subordinations for Functions with Positive Real Part, Complex Anal. Oper. Theory. **12** (2018), no.5, 1179–1191.
[6] W. C. Ma and D. Minda, A unified treatment of some special classes of univalent functions, in *Proceedings of the Conference on Complex Analysis (Tianjin, 1992)*, 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge, MA.

[7] S. S. Miller and P. T. Mocanu, Differential subordinations and univalent functions, Michigan Math. J. **28** (1981), no. 2, 157–172.
[8] S. S. Miller and P. T. Mocanu, *Differential subordinations*, Monographs and Textbooks in Pure and Applied Mathematics, 225, Marcel Dekker, Inc., New York, 2000.
[9] A. Naz, S. Nagpal and V. Ravichandran, Star-likeness associated with the exponential function, Turkish J. Math. **43** (2019), no. 3, 1353–1371.
[10] M. Nunokawa and J. Sokół, An improvement of Ozaki's condition, Appl. Math. Comput. **219** (2013), no. 22, 10768–10776.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: shagun05banga@gmail.com

Digital Signal Processing 109 (2021) 102909

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

On the performance of DF-based multi-hop system over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-extreme fading channels

Tau Raphael Rasethuntsa [a], Manpreet Kaur [b], Sandeep Kumar [c,*], Puspraj Singh Chauhan [d], Kuldeep Singh [e]

[a] Mathematical Engineering, Istanbul Technical University, Istanbul, Turkey
[b] Department of Computer Science & Engineering, Delhi Technological University, Delhi, India
[c] Central Research Laboratory, Bharat Electronics Limited, Ghaziabad, Uttar Pradesh, India
[d] Department of Electronics & Comm. Engg., M. M. M. University of Technology, Gorakhpur, Uttar Pradesh, India
[e] Department of Electronics & Comm. Engg. Malaviya National Institute of Technology, Jaipur, India

ARTICLE INFO

Article history:
Available online 24 November 2020

Keywords:
DF relaying
Ergodic capacity
Marcum Q-function
Nuttall Q-function
Decoding error probability

ABSTRACT

Several emerging application scenarios for the multi-hop wireless networking have been introduced in 5G and beyond networks, to provide ultra-low latency and much higher data rate to the end users. In this work, an integrated performance evaluation of a decode-and-forward (DF) multi-hop wireless communication system is undertaken over the non-linear generalized $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading models. Novel closed-form expressions for the amount of fading (AoF), channel capacity under various adaptive transmission techniques are derived. Additionally, decoding error probability for short packet communication is derived and analyzed. The provided expressions are very useful in designing the channel/link quality for practical multi-hop communication networks. The effects of the fading parameters and the number of hops on the system performance are demonstrated. The correctness of the proposed solutions has been corroborated by comparing them with Monte Carlo simulation results.
© 2020 Elsevier Inc. All rights reserved.

1. Introduction

While wireless communication will be the core technique, a direct communication between the transmitter and the receiver is faced with many limitations. In particular, communication over long distances is either possible through prohibitively high transmission power or using multi-hope relaying techniques. The use of multi-hope relaying increases the network coverage area with high achievable data rate. Several emerging application scenarios for the multi-hop wireless networking have been found in 5G and beyond technologies. For example, in device-to-device communication, the transmitted information reaches its final destination via several intermediate devices, known as relays [1]. DF-based multi-hop systems have several other applications in ad-hoc networks, sensor networks and microwave links among many others [2]. The main idea behind multi-hope communication is that the channel from a source node to a destination node is split into multiple possibly shorter links by using some intermediate nodes. Based on the nature and complexity of the relaying system, multi-hop communication systems can be classified as (i) decode-and-forward (DF) and (ii) amplify-and-forward [1]. For a DF-based multi-hop system, the relaying node receives the encoded signal, decodes it to regenerate the original symbol and forwards the newly encoded signal to the next node [2]. The modeling of mobile radio channel is a challenging task as the wireless signal is subjected to various hurdles/corruptions in its propagation path from the transmitter to the receiver [3]. Accurate characterization of the wireless channel is therefore of paramount importance for the realistic assessment of wireless system performance.

Several statistical generalized distributions like $\alpha-\eta-\kappa-\mu$, $\alpha-\eta-\mu$, $\alpha-\kappa-\mu$, $\alpha-\mu$, $\eta-\mu$ and $\kappa-\mu$ etc. are developed in the literature to model small-scale fluctuations of the fading channel envelope. Recently, the effective capacity analysis over $\alpha-\eta-\kappa-\mu$ fading channel has been proposed as a generalized solution for almost all of the multi-path fading channels [4]. The secrecy performance over correlated $\alpha-\mu$ distribution has been proposed in [5], while the channel capacity and energy detection performance over $\alpha-\eta-\mu$ distribution have been presented in [6] and [7]. The $\kappa-\mu$ and $\alpha-\kappa-\mu$ are the widely accepted models, which accounts for the signal propagation in line-of-sight (LOS) communications. In [8], the error probability for an uplink

* Corresponding author.
E-mail addresses: tauzand@gmail.com (T.R. Rasethuntsa), manpreettiet@gmail.com (M. Kaur), sann.kaushik@gmail.com (S. Kumar), puspraj.chauhan@gmail.com (P.S. Chauhan), kuldeep.ece@mnit.ac.in (K. Singh).

https://doi.org/10.1016/j.dsp.2020.102909
1051-2004/© 2020 Elsevier Inc. All rights reserved.

communication over non-homogeneous $\kappa-\mu$ generalized fading channels is investigated. Different performance metrics over shadowed $\kappa-\mu$ composite fading channel is discussed in [9], [10] and [11]. Motivated by the pioneering works in [12] and [13] the generalized $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading models were proposed in [14] and [15]. The importance of these models lies in the fact, that these channels can accommodate most of the practical fading models as special cases. In particular, when $\alpha=2$, the $\alpha-\kappa-\mu$ distribution reduces to the $\kappa-\mu$ distribution [12] and for $\kappa\to 0$, the $\alpha-\mu$ is obtained [16]. From the $\kappa-\mu$ and $\alpha-\mu$, models, one can obtain other distributions including the Nakagami-m, Weibull, Rice, one-sided Gaussian and Rayleigh models for particular values of α, κ or μ. In [17], the performance of a dual-branch selection combining (SC) over $\alpha-\kappa-\mu$ fading channels is evaluated. The spectrum sensing capability of a cooperative system over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels is investigated in [18]. The work in [19] and [20] have studied the energy detector performance over generalized shadowed composite fading distributions. The $\alpha-\kappa-\mu$-Extreme fading model is derived from the $\alpha-\kappa-\mu$ distribution by allowing $\mu\to 0$ and $\kappa\to\infty$ in order to better characterize small-scale variations of mobile radio propagation under non-linear severe multi-path fading effects usually encountered in enclosed environments [15]. For $\alpha=2$, the $\alpha-\kappa-\mu$-Extreme distribution reduces to a linear severe fading model, the $\kappa-\mu$-Extreme distribution [13].

The performance of DF-based relaying wireless systems and networks has been extensively investigated under several fading distributions. The channel capacity of a multi-hop relaying system over Rayleigh fading channels has been studied in [21] whereas the end-to-end performance of a DF relaying system over Ricean and Nakagami-m fading channels is explored in [22]. The authors in [23] investigated the bit-error-rate performance of a DF multi-hop cognitive networks over Rayleigh fading channel. The performance of multi-hop relaying in industrial wireless sensor networks over Nakagami-m fading channels is presented in [24]. The work in [25] has studied the system performance of multi-hop relay network in a quasi-static slow fading environment. Recently the outage probability and bit error rate performance of a DF-based multi-hop system over the generalized fading channels has been performed in [26]. Although there are many works related to the study of DF-based multi-hop relaying systems over different fading channels. None of the previous works has investigated the end-to-end channel capacity and decoding error probability over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels, which represent our key contribution. It is important to study the performance metrics of the multi-hop wireless systems over the generalized fading models. Such investigation is essential in order to implement practical wireless system realizations with predefined performance levels. Motivated by this void, we have derived the closed-form mathematical expressions for amount of fading (AoF), channel capacity under various adaptive transmission schemes and decoding error probability for short packet communication for end-to-end DF relaying system over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels. The expressions derived here are presented in a simple manner to optimize clarity and readability. Furthermore, the results produced here are generalized expressions and valid for other fading distributions as special cases.

The rest of the paper is organized as follows: The system and channel model are described in Section 2. In Section 3, we derived the closed-form expressions of performance metrics like AoF, decoding error probability and channel capacity under different transmission policy for the DF-based multi-hop wireless system over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels. The asymptotic series expansion of decoding error probability is presented in Section 4. Performance evaluation results and discussions are provided in Section 5. Finally, we round up the paper in Section 6 with conclusions and possible future considerations.

Table 1
Notations.

S	Source node
D	Destination node
γ	Signal-to-noise-ratio
$p_\gamma(\gamma)$	Probability density function (PDF)
$\delta(\cdot)$	Dirac delta function
$Q(\cdot,\cdot)$	Generalized Marcum Q-function
$I(\cdot)$	Modified Bessel functions of first kind
$\bar{\gamma}$	Average received signal-to-noise-ratio
α	Severity of the channel non-linearity
κ	Ratio of the power of dominant components to the scattered waves
μ	Number of the multi-path cluster
$P_\gamma(\gamma)$	Cumulative distribution function (CDF)
$\mathbb{E}(.)$	Expectation operator
$\tilde{\Phi}(.;.;.)$	Regularized confluent hypergeometric function
$\Gamma(.)$	Gamma function
$e^{(.)}$	Exponential function
AoF	Amount of fading
$Var(.)$	Variance operator
$G(\cdot)$	Meijer's functions
R	Achievable rate
N	Block length
ϵ	Decoding error probability
$Q^{-1}(\cdot)$	Inverse Q-function
V	Channel dispersion
$\bar{\epsilon}$	Average decoding error probability
$\mathcal{Q}(\cdot,\cdot)$	Nuttall Q-function
$\Gamma(a,x)$	Incomplete gamma function
B	Bandwidth of the channel
$\ln(.)$	Natural logarithm
$\log_z(.)$	Logarithm function to the base z
$\mathcal{C}_C$	CIFR channel capacity
$\mathcal{C}_T$	TIFR channel capacity
$\mathcal{C}_O$	ORA channel capacity
$H(\cdot)$	Fox-H function
$\mathcal{C}_P$	OPRA channel capacity
γ_0	Optimum cut-off signal-to-noise-ratio
$\mathcal{R}$	Spectrum efficiency

2. System and channel models

We consider a wireless relay network in which a source node S transmits its data packets to the destination node D through $n-1$ relay devices operating over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels. The $n+1$ nodes (S, D, and the intermediate relay) are assumed to be sufficiently and equally spaced. Hence, a device can only receive the transmitted data of its previous-hop adjacent node. Intermediate terminals always perform hard decisions on the received symbols before forwarding them to their respective successor node. We assume that every relay processes only the signals received from its preceding nodes, which allows for reducing the computational costs and hardware complexity at the receiver of each node. Furthermore, it is appropriate for wireless sensor networks, which require fixed processing complexity at each node. The analysis over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channel is performed under Case I and Case II respectively. All the symbols and notations used in this work have been defined in Table 1.

2.1. PDF and CDF of end-to-end SNR

The PDF and CDF expressions for the end-to-end SNR of DF relaying system over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channel are given by [26] as

Case I.

$$p_\gamma(\gamma) = \mathcal{A}\delta(\gamma) + \mathcal{K}_n\bar{\mathcal{A}}\frac{\gamma^{(\frac{\alpha_n}{2})\omega_n-1}}{e^{\sigma_n\gamma^{\frac{\alpha_n}{2}}}} I_{\mu_n-1}\left(\sqrt{\theta_n\gamma^{\frac{\alpha_n}{2}}}\right), \tag{1}$$

where $\mathcal{A} = 1 - \prod_{k=1}^{n-1} Q_{\mu_k}\left(\sqrt{2\kappa_k\mu_k}, \sqrt{2\sigma_k\gamma^{\frac{\alpha_k}{2}}}\right)$ is the outage probability of the first $n-1$ hops.

$$P_\gamma(\gamma) = 1 - \bar{\mathcal{A}}Q_{\mu_n}\left(\sqrt{2\kappa_n\mu_n}, \sqrt{2\sigma_n\gamma^{\frac{\alpha_n}{2}}}\right). \tag{2}$$

To improve clarity in presentation, we have set

$$\mathcal{K}_n = \frac{\alpha_n\sigma_n^{\omega_n}e^{-\mu_n\kappa_n}}{2(\mu_n\kappa_n)^{\omega_n-1}},\ \omega_n = \frac{\mu_n+1}{2},\ \sigma_n = \frac{\mu_n(\kappa_n+1)}{\bar{\gamma}_n^{\frac{\alpha_n}{2}}},$$
$$\theta_n = 4\mu_n\kappa_n\sigma_n, \tag{3}$$

with $\alpha_n, \mu_n, \kappa_n > 0$.

Case II.

$$p_\gamma^e(\gamma) = \mathcal{A}^e\delta(\gamma) + \bar{\mathcal{A}}^e\left[\frac{a_n\gamma^{\frac{\alpha_n}{4}-1}}{e^{b_n\gamma^{\frac{\alpha_n}{2}}}} I_1\left(\sqrt{c_n\gamma^{\frac{\alpha_n}{2}}}\right) + \delta^*(\gamma)\right], \tag{4}$$

where $\delta^*(\gamma) = \frac{e^{-2m_n}}{2\sqrt{\gamma\bar{\gamma}_n}}\delta\left(\sqrt{\frac{\gamma_n}{\bar{\gamma}}}\right)$ and

$$\mathcal{A}^e = 1 - \prod_{k=1}^{n-1} Q_0\left(2\sqrt{m_k}, 2\sqrt{m_k(\gamma_k/\bar{\gamma}_k)^{\frac{\alpha_k}{2}}}\right).$$

$$P_\gamma^e(\gamma) = 1 - \bar{\mathcal{A}}^e Q_0\left(2\sqrt{m_n}, 2\sqrt{m_n(\gamma/\bar{\gamma})^{\frac{\alpha_n}{2}}}\right), \tag{5}$$

where

$$a_n = \frac{\alpha_n m_n e^{-2m_n}}{\bar{\gamma}_n^{\alpha_n/4}}, b = \frac{2m_n}{\bar{\gamma}_n^{\frac{\alpha_n}{2}}}, c_n = \frac{(16m_n^2)}{\bar{\gamma}_n^{\frac{\alpha_n}{2}}}, \tag{6}$$

with $\alpha_n, \mu_n, \kappa_n > 0$ and $m_n = \mu_n(\kappa_n+1)^2/(2\kappa_n+1)$ as the Nakagami parameter which is inversely proportional to the fading severity.

2.2. Moments of end-to-end SNR

Generalized moments of the end-to-end SNR are useful for obtaining rth-order performance metrics such as the mean, variance as well as kurtosis and skewness of the distribution of γ. The rth moment of a positive random variable X with density f_X is defined by

$$\mathbb{E}\langle X^r\rangle \triangleq \int_0^\infty x^r f_X(x)dx. \tag{7}$$

For Case I, substituting (1) into (7) and setting $y = \gamma^{\alpha_2/2}$, the rth moment of γ can be written with the aid of [27, Eq. 6.643(2)] and [27, Eq. 6.220(2)] as

$$\mathbb{E}\langle\gamma^r\rangle = \frac{\bar{\mathcal{A}}\bar{\gamma}_n^r\Gamma\left(\frac{2r}{\alpha_n}+\mu_n\right)\tilde{\Phi}\left(\frac{2r}{\alpha_n}+\mu_n;\mu_n;\mu_n\kappa_n\right)}{e^{\mu_n\kappa_n}\left[\mu_n(\kappa_n+1)\right]^{\frac{2r}{\alpha_n}}}. \tag{8}$$

Remark 1. Moments of the special cases:

a) In case of Rice distribution, (i.e. when $n=1$, $\alpha_1=2$ and $\mu_1=1$) (8) reduces to [28, Eq. 2.18] after the application of [27, Eq. 9.212(1)] as

$$\mathbb{E}\langle\gamma^r\rangle = \frac{\bar{\gamma}_1^r\Gamma(r+1)\tilde{\Phi}(-r;1;-\kappa_1)}{(\kappa_1+1)^r}. \tag{9}$$

b) When $n=1$, $\alpha_1=2$ and $\kappa_1\to 0$ (Nakagami-m case) we note that [27, Eq. 9.210(1)] implies that $\tilde{\Phi}\left(\frac{2r}{\alpha_1}+\mu_1;\mu_1;\mu_1\kappa_1\right)\to 1$ as $\kappa_1\to 0$. Consequently, (8) reduces to [28, Eq. 2.23] as

$$\mathbb{E}\langle\gamma^r\rangle = \frac{\bar{\gamma}_1^r\Gamma(\mu_1+r)}{\Gamma(\mu_1)\mu_1^r}. \tag{10}$$

c) $\mathbb{E}\langle\gamma^r\rangle$ for the Rayleigh case can be obtained by setting $\mu_1=1$ in (10). This corroborates the correctness of (8).

Similarly, for Case II, we have the rth moment of γ as

$$\mathbb{E}^e\langle\gamma^r\rangle = \frac{\bar{\mathcal{A}}^e\bar{\gamma}_n^r\Gamma\left(\frac{2r}{\alpha_n}+1\right)\tilde{\Phi}\left(\frac{2r}{\alpha_n}+1;2;2m_n\right)}{e^{2m_n}(2m_n)^{\frac{2r}{\alpha_n}-1}}. \tag{11}$$

It is also worth noting that as expected, upon substituting $r=1$ in (8) and (11), one obtains $\bar{\gamma} = \bar{\mathcal{A}}\bar{\gamma}_n$ and $\bar{\gamma} = \bar{\mathcal{A}}^e\bar{\gamma}_n$ for each case, respectively

3. Performance analysis of multi-hop DF-based relay systems

3.1. Amount of fading

The second-order AoF is a useful performance metric that can be used to account for the severity of fading. AoF is defined as the ratio of the variance to the square average SNR per symbol, $AoF_\gamma^{(2)} \triangleq \mathrm{Var}(\gamma)/\bar{\gamma}^2$. That is,

$$AoF_\gamma^{(2)} \triangleq \frac{\mathbb{E}\langle\gamma^2\rangle - \mathbb{E}\langle\gamma\rangle^2}{\mathbb{E}\langle\gamma\rangle^2}. \tag{12}$$

This can be generalized to higher-order fading as $AoF_\gamma^{(n)} \triangleq \mathbb{E}\langle\gamma^n\rangle/(\mathbb{E}\langle\gamma\rangle^n) - 1$. Substitution of $\mathbb{E}\langle\gamma\rangle$ and $\mathbb{E}\langle\gamma^2\rangle$ from (8) into (12) yields after some basic simplifications the second-order AoF of γ under Case I as

$$AoF_\gamma = \frac{e^{\mu_n\kappa_n}\Gamma\left(\frac{4}{\alpha_n}+\mu_n\right)\tilde{\Phi}\left(\mu_n+\frac{4}{\alpha_n};\mu_n;\kappa_n\mu_n\right)}{\bar{\mathcal{A}}\Gamma\left(\frac{2}{\alpha_n}+\mu_n\right)^2\tilde{\Phi}\left(\mu_n+\frac{2}{\alpha_n};\mu_n;\kappa_n\mu_n\right)^2} - 1. \tag{13}$$

Likewise, for Case II, the second-order amount of fading can be derived using (11) in (12) resulting in

$$AoF_\gamma^e = \frac{e^{2m_n}\Gamma\left(\frac{4}{\alpha_n}+1\right)\tilde{\Phi}\left(\frac{4}{\alpha_n}+1;2;2m_n\right)}{\bar{\mathcal{A}}^e 2m_n\Gamma\left(\frac{2}{\alpha_n}+1\right)^2\tilde{\Phi}\left(\frac{2}{\alpha_n}+1;2;2m_n\right)^2} - 1. \tag{14}$$

3.2. Decoding error probability for short-packet communications

In 5G wireless networks, short packet communication is used to support low latency in real-time functionality. In short-packet communication, the decoding error probability at the receiver is not negligible since the length of the block is finite. This is different from the Shannon capacity theorem, in which the decoding error probability becomes negligible as the block length approaches infinity. For short-packet communication, the achievable rate with finite block length and for a given decoding error probability and the ith hop, $i = 1, \ldots, n$, can be expressed using [29] as

$$R = C(\gamma) - \sqrt{\frac{V(\gamma)}{N}} Q^{-1}(\epsilon), \tag{15}$$

where $C(\gamma) = \log_2(1+\gamma)$ is the Shannon's capacity and the channel dispersion given by $V(\gamma) = \frac{\gamma(\gamma+2)}{2(\gamma+1)^2} (\log_2(e))^2$. Using (15), the average decoding error probability of the ith hop for the fading channel can be given as

$$\bar{\epsilon} = \int_0^{\infty} Q\left(\frac{C(\gamma) - R}{\sqrt{V(\gamma)/N}}\right) p_\gamma(\gamma) d\gamma, \tag{16}$$

The integration in (16) is very complex and challenging to compute in closed form. Two methods are used to find the approximate solution of (16).

Method I. Using linear approximation of the term with the Q-function [30], we have

$$Q\left(\frac{C(\gamma) - R}{\sqrt{V(\gamma)/N}}\right) \approx \begin{cases} 1 & \gamma \le \gamma^- \\ \frac{1}{2} - \lambda\sqrt{N}(\gamma - \gamma_{th}) & \gamma^- < \gamma < \gamma^+ \\ 0 & \gamma \ge \gamma^+ \end{cases}, \tag{17}$$

where

$$\lambda = \frac{1}{2\pi\sqrt{2^{2R}-1}}, \gamma_{th} = 2^R - 1 \text{ and } \gamma^{\pm} = \left(\gamma_{th} \pm \frac{1}{2\lambda\sqrt{N}}\right). \tag{18}$$

Applying the approximation of (17) in (16) and using simple mathematical manipulation, the average decoding probability for the ith hop of the DF system experiencing $\alpha - \kappa - \mu$ fading can be calculated as

$$\begin{aligned} \bar{\epsilon} &= \int_0^{\gamma^-} p_\gamma(\gamma) d\gamma + \int_{\gamma^-}^{\gamma^+} \left[\frac{1}{2} - \lambda\sqrt{N}(\gamma - \gamma_{th})\right] p_\gamma(\gamma) d\gamma \\ &= P_\gamma(\gamma^-) + \left(\frac{1}{2} + \lambda\gamma_{th}\sqrt{N}\right)\left[P_\gamma(\gamma^+) - P_\gamma(\gamma^-)\right] \\ &\quad - \lambda\sqrt{N} \underbrace{\int_{\gamma^-}^{\gamma^+} \gamma\, p_\gamma(\gamma) d\gamma}_{\mathcal{J}}, \end{aligned} \tag{19}$$

The integral $\mathcal{J}$ in (19) can be re-written as

$$\mathcal{J} = \int_0^{\gamma^+} \gamma\, p_\gamma(\gamma) d\gamma - \int_0^{\gamma^-} \gamma\, p_\gamma(\gamma) d\gamma. \tag{20}$$

Therefore, using (1) and [31, Eq. 2], it follows after some basic manipulations that

$$\begin{aligned} \int_0^{x} \gamma\, p_\gamma(\gamma) d\gamma &= 1 - \bar{A}\frac{2\sigma_n^{\omega_n} (2\sigma_n)^{\omega_n + 2/\alpha_n - 1}}{(\mu_n\kappa_n)^{\omega_n - 1}} \\ &\quad \times \mathcal{Q}_{2\omega^* - 1, \mu_n - 1}\left(\sqrt{\frac{\theta_n}{2\sigma_n}}, \sqrt{2\sigma_n (x)^{\frac{\alpha_n}{2}}}\right), \end{aligned} \tag{21}$$

where the Nuttall Q-function has parameters $\mu_n > 1$ and $\omega^* = \omega_n + 2/\alpha_n > 1$. Since the Nuttall Q-function has no implementation in standard packages, its series representation has been proposed in [32, Eq. 7] as

$$\mathcal{Q}_{v,w}(a,b) = \sum_{k=0}^{\infty} \frac{a^{2k} e^{-\frac{a^2}{2}} \Gamma\left(\frac{v+w+2k+1}{2}, \frac{b^2}{2}\right)}{k!\Gamma(v+k+1) 2^{\frac{v-w+2k+1}{2}}}. \tag{22}$$

The form in (22) can be used for numerical computations.

Method II. It is shown in [33] that the low SNR approximation of the average decoding probability for the ith hop of the DF system is given by

$$\bar{\epsilon} \approx P_r(r_o), \quad r_o = 2^R - 1, \tag{23}$$

where P_r is the CDF of the fading channel. Therefore, putting (2) and (5) in (23) will yield the desired expressions for the $\alpha - \kappa - \mu$ and $\alpha - \kappa - \mu$-Extreme fading channels respectively.

3.3. Channel capacity

Tremendous growth in the field of wireless communication compelled researchers to determine the capacity limit i.e., the maximum data rate that can be sustained by a channel, without putting a limit on the complexity or delay of the encoder and decoder. Many authors have already derived the mathematical expression of channel capacity over different fading channel. There are different policies for transmission that depends on the availability of channel state information (CSI) at transmitter and the receiver. In this section, we investigate the capacity of $\alpha - \kappa - \mu$ and $\alpha - \kappa - \mu$-Extreme fading channels under various adaptive transmission schemes. This is an important topic which has been covered in works such as [34]. All the results derived in this section are presented in closed-form.

3.3.1. Optimal rate adaptation

When the average transmitted power remains constant, usually as a result of CSI being available only at the receiver side such scheme is known as ORA. The channel capacity of the received end-to-end SNR γ with density function $p_\gamma(\gamma)$ under ORA (also known as ergodic capacity) for Case I and Case II is derived in Appendix A and the final results are given in Table 2.

Table 2
Ora channel capacity.

$$\mathcal{C}_O = \frac{B\bar{A}\pi\alpha_n\mathcal{K}_n}{2n\sigma_n^{\omega_n}\ln 2} H^{0,1:1,2;1,0}_{1,0:2,2;1,3}\left(\begin{array}{c|c} \left(\frac{1}{\sigma_n}\right)^{\frac{2}{\alpha_n}} & \mathbf{\Omega}(\alpha_n, \mu_n, \omega_n) \\ \frac{\theta_n}{4\sigma_n} & \mathbf{\Pi}(\alpha_n, \mu_n, \omega_n) \end{array}\right). \tag{24}$$

$$\mathcal{C}_O^e = \frac{B\bar{A}^e\pi\alpha_n a_n}{2n\sqrt{b_n}\ln 2} H^{0,1:1,2;1,0}_{1,0:2,2;1,3}\left(\begin{array}{c|c} \left(\frac{1}{b_n}\right)^{\frac{2}{\alpha_n}} & \mathbf{\Omega}\left(\alpha_n, 2, \frac{1}{2}\right) \\ \frac{c_n}{4b_n} & \mathbf{\Pi}\left(\alpha_n, 2, \frac{1}{2}\right) \end{array}\right). \tag{25}$$

$$\mathbf{\Omega} \equiv \left(1 - \omega_n; \frac{2}{\alpha_n}, 1\right) : (1,1), (1,1); \left(\frac{\mu_n}{2}, 1\right).$$

$$\mathbf{\Pi} \equiv - : (1,1), (0,1); (\omega_n - 1, 1), (1 - \omega_n, 1), \left(\frac{\mu_n}{2}, 1\right).$$

3.3.2. Optimal power and rate adaptation

When both the receiver and the transmitter have the information of the CSI, and the transmitter is capable of adapting the transmitted power constantly to maximize the capacity, then such a policy is kept under the OPRA scheme. The intention behind this adaptation is to assign more power to the channel with high SNR value and lesser power to the channel with lower SNR. The OPRA technique achieves the highest possible capacity with CSI by employing a multiplexed multiple code-book design in order to match

the transmission power and rate of the system [35]. The OPRA capacity in DF-based multi-hop systems for both Case I and Case II is derived in Appendix B as shown in Table 3.

Table 3
OPRA channel capacity.

$$\mathcal{C}_P = \frac{4\gamma_o^{\frac{\alpha_n}{2}\omega_n}\pi B\mathcal{K}_n}{\alpha_n^2 n \ln 2 \bar{\mathcal{A}}^{-1}} G^{0,2:1,0;1,0}_{2,2:0,1;1,3}\left(\begin{matrix}\sigma_n\gamma_o^{\frac{\alpha_n}{2}} \\ \frac{\theta_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\end{matrix}\middle|\begin{matrix}\spadesuit(\mu_n,\omega_n) \\ \clubsuit(\mu_n,\omega_n)\end{matrix}\right). \tag{26}$$

$$\mathcal{C}_P^e = \frac{4\pi\bar{\mathcal{A}}^e\gamma_o^{\frac{\alpha_n}{4}}Ba_n}{\alpha_n^2 n \ln 2} G^{0,2:1,0;1,0}_{2,2:0,1;1,3}\left(\begin{matrix}b_n\gamma_o^{\frac{\alpha_n}{2}} \\ \frac{c_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\end{matrix}\middle|\begin{matrix}\spadesuit\left(2,\frac{1}{2}\right) \\ \clubsuit\left(2,\frac{1}{2}\right)\end{matrix}\right). \tag{27}$$

$\spadesuit \equiv (1-\omega_n;1,1),(1-\omega_n;1,1):-;\left(\frac{\mu_n}{2},1\right).$

$\clubsuit \equiv (-\omega_n;1,1);(-\omega_n;1,1):(0,1);(\omega_n-1,1),(1-\omega_n,1),\left(\frac{\mu_n}{2},1\right).$

3.3.3. Channel inversion with fixed rate

Another simple technique which offers easy implementation is to maintain a constant SNR at the receiver with fixed-rate modulation and fixed code design, Simon and Alouini [28] is known as CIFR policy. This policy is also known as zero-outage capacity as it maintains the fixed data rate regardless of channel condition with zero channel outage. Under this policy, the transmitter allocates higher power to the corresponding channel encountering low SNR value so that the non-fluctuating received power is maintained. Under this technique, the channel capacity of the end-to-end SNR of a n-link DF-based multi-hop system with density $p_\gamma(\gamma)$ is given by

$$\mathcal{C}_c \triangleq \frac{B}{n}\log_2\left(1+\frac{1}{\int_0^\infty \left(p_\gamma(\gamma)/\gamma\right)d\gamma}\right). \tag{28}$$

Since the expression obtained using (8) and (11) are highly restrictive, we can obtain more general ones by employing a similar procedure to the one used in deriving the MGF function in [26]. We only express the Bessel function in each SNR PDF in terms of $G(\cdot)$ and then reverse the order of integration and after some simplifications the single Mellin-Barnes integral is obtained as

$$\int_0^\infty \frac{p_\gamma(\gamma)}{\gamma}d\gamma = \frac{2\pi\mathcal{K}_n\bar{\mathcal{A}}}{\alpha_n\sigma_n^{\omega_n-\frac{2}{\alpha_n}}}\frac{1}{2\pi i}\oint_{L_r}\frac{\Gamma\left(\omega_n-\frac{2}{\alpha_n}+r\right)}{\Gamma(\omega_n+r)} \times \frac{\Gamma(\omega_n-1-r)\left(\frac{\theta_n}{4\sigma_n}\right)^r dr}{\Gamma\left(1-\frac{\mu_n}{2}+r\right)\Gamma\left(\frac{\mu_n}{2}-r\right)}, \tag{29}$$

where $Re\left(\omega_n-\frac{2}{\alpha_n}+r\right)>0$. Comparing the contour integral in (29) with [36, Eq. 1.1] yields the CIFR channel capacity for Case I as

$$\frac{n\mathcal{C}_c}{B} = \log_2\left\{1+\left[\frac{2\pi\mathcal{K}_n\bar{\mathcal{A}}}{\alpha_n\sigma_n^{\omega_n-\frac{2}{\alpha_n}}}G^{1,1}_{2,3}\left(\frac{\theta_n}{4\sigma_n}\middle|\begin{matrix}\blacktriangle \\ \blacktriangledown\end{matrix}\right)\right]^{-1}\right\}, \tag{30}$$

where $\blacktriangle \equiv 1+\frac{2}{\alpha_n}-\omega_n, \frac{\mu_n}{2}$ and $\blacktriangledown \equiv \omega_n-1, 1-\omega_n, \frac{\mu_n}{2}$. In an analogous fashion, the CIFR channel capacity for Case II can be derived with the contour condition $Re\left(1/2-\frac{2}{\alpha_n}+r\right)>0$ as

$$\frac{n\mathcal{C}_c^e}{B} = \log_2\left\{1+\left[\frac{2\pi a_n\bar{\mathcal{A}}^e}{\alpha_n b_n^{\frac{1}{2}-\frac{2}{\alpha_n}}}G^{1,1}_{2,3}\left(\frac{c_n}{4b_n}\middle|\begin{matrix}\blacktriangle^* \\ \blacktriangledown^*\end{matrix}\right)\right]^{-1}\right\}, \tag{31}$$

where $\blacktriangle^* \equiv \frac{1}{2}+\frac{2}{\alpha_n}, 1$ and $\blacktriangledown^* \equiv -\frac{1}{2}, \frac{1}{2}, 1$.

3.3.4. Truncated channel inversion with fixed rate

In severe fading conditions, the channel inversion technique with fixed rate may suffer large capacity penalties as compared to other techniques. To circumvent this shortcoming, a modified inversion of $\mathcal{C}_T$ which inverts the channel fading above γ_o is $\mathcal{C}_c$ which is defined as

$$\mathcal{C}_T \triangleq \frac{B}{n}\log_2\left(1+\frac{1}{\int_{\gamma_o}^\infty p_\gamma(\gamma)/\gamma d\gamma}\right)Pr\left(\gamma>\gamma_o\right). \tag{32}$$

Note that $\mathcal{C}_T \to \mathcal{C}_c$ from above as $\gamma_o \to 0$. That is, the gap between $\mathcal{C}_T$ and $\mathcal{C}_c$ depends on $Pr\left(\gamma>\gamma_o\right)$ in (32). Making the substitution $y=\gamma/\gamma_o$, the integral in (32) can be re-written as

$$\mathcal{J} = \int_{\gamma_o}^\infty p_\gamma(\gamma)/\gamma d\gamma = \int_1^\infty \frac{1}{y}f_y(\gamma_o y)dy, \tag{33}$$

Substituting (1) into (33) and following the simple procedure undertaken in deriving OPRA channel capacity yields (34) and (35) as shown in Table 4

Table 4
TIFR channel capacity integral $\mathcal{J}$.

$$\mathcal{J} = \frac{-2\pi\bar{\mathcal{A}}\left(\gamma_o^{\frac{\alpha_n}{2}\omega_n}\mathcal{K}_n\right)}{\alpha_n\gamma_o}G^{0,1:1,0;1,0}_{1,2:0,1;1,3}\left(\begin{matrix}\sigma_n\gamma_o^{\frac{\alpha_n}{2}} \\ \frac{\theta_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\end{matrix}\middle|\begin{matrix}\spadesuit^*(\mu_n,\omega_n) \\ \clubsuit^*(\mu_n,\omega_n)\end{matrix}\right). \tag{34}$$

$$\mathcal{J} = \frac{-2\pi\bar{\mathcal{A}}^e\left(\gamma_o^{\frac{\alpha_n}{4}}a_n\right)}{\alpha_n}G^{0,1:1,0;1,0}_{1,1:0,1;1,3}\left(\begin{matrix}b_n\gamma_o^{\frac{\alpha_n}{2}} \\ \frac{c_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\end{matrix}\middle|\begin{matrix}\spadesuit^*\left(2,\frac{1}{2}\right) \\ \clubsuit^*\left(2,\frac{1}{2}\right)\end{matrix}\right). \tag{35}$$

$\spadesuit^* \equiv \left(1+\frac{2}{\alpha_n}-\omega_n;1,1\right):-;\left(\frac{\mu_n}{2},1\right).$

$\clubsuit^* \equiv \left(\frac{2}{\alpha_n}-\omega_n;1,1\right),\left(\frac{2}{\alpha_n}-\omega_n;1,1\right):(0,1);(\omega_n-1,1),(1-\omega_n,1),\left(\frac{\mu_n}{2},1\right).$

The double Mellin-Barnes contour integral must satisfy $Re\left(\omega_n-\frac{2}{\alpha_n}+r+t\right)<0$. Therefore, putting (34) and (35) into (32) gives the TIRF channel capacity for Case I and Case II, respectively. Alternatively, for Case I, another closed-form expression for the integral in (32) can be obtained by making use of the consecutive substitutions $x=\gamma^{\alpha_n/2}$ followed by $\sqrt{x}=u/\sqrt{2\sigma_n}$ which yield

$$\mathcal{J} = \frac{\bar{\mathcal{A}}(2\sigma_n)^{\frac{2}{\alpha_n}}}{2(2\mu_n\kappa_n)^{\omega_n-1}}\mathcal{Q}_{\mu^*,\mu_n-1}\left(\sqrt{2\mu_n\kappa_n},\sqrt{2\sigma_n\gamma_o^{\frac{\alpha_n}{2}}}\right), \tag{36}$$

where $\mu^*=\mu_n-\frac{4}{\alpha_n}$, $\alpha_n\mu_n>4$.

$$\mathcal{Q}_{M,N}(a,b) = \int_b^\infty t^M\exp\left(-\frac{t^2+a^2}{2}\right)I_N(at)dt, \tag{37}$$

with $b, M, N \geq 0$ and $a>0$. A similar procedure can be followed for Case II. However, the resulting integral does not satisfy the $M\geq 0$ condition and hence fails to attain a Nuttall Q-function representation. The Nuttall Q-function is a generalization of the Marcum Q-function. Specifically, when $M=N+1$, $\mathcal{Q}_{N+1,N}(a,b)=a^N Q_{N+1}(a,b)$. The latter is clearly evident from (36) in the case

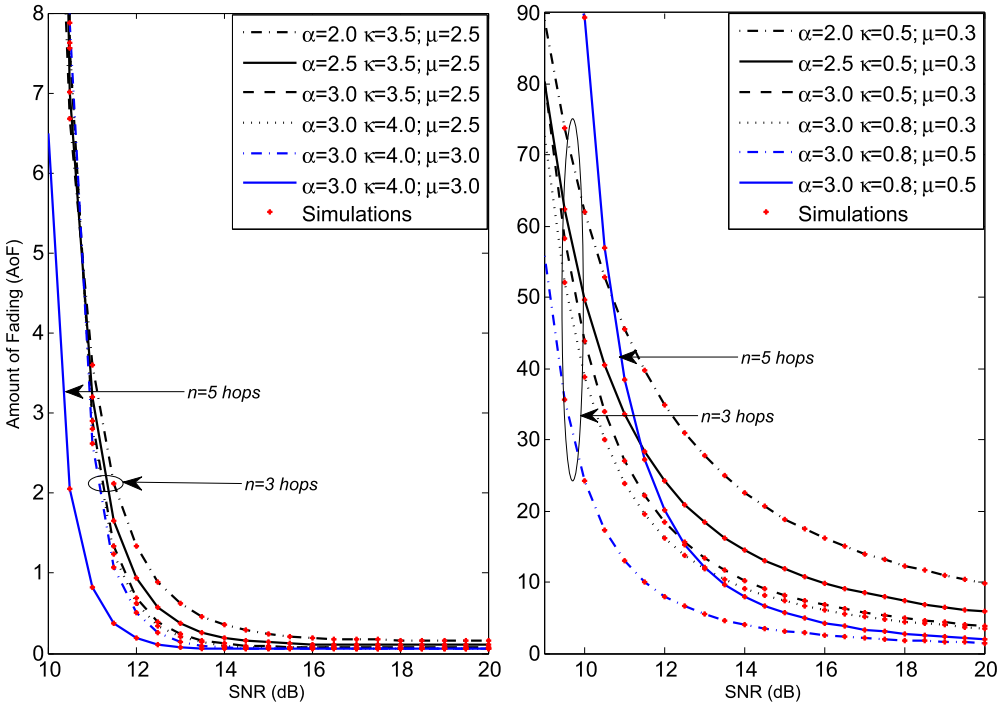


Fig. 1. AoF curves for the $\alpha-\kappa-\mu$ case showing the effects of variation in model parameter values and hop numbers under moderate and severe fading conditions. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$\alpha = 2$. There are currently no implementations of the Nuttall Q-function in standard packages. Nonetheless, the computation of (37) can be carried out via numerical integration.

4. Poincare asymptotic series expansion

In this section, the Poincare asymptotic series expansion of the decoding error probability for both the $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading models shall be discussed. The PDFs of the SNRs for the nth hop of Case I and Case II fading channels have Poincare asymptotic series expansion near the origin and is given by [26] as

$$p_\gamma(\gamma) \doteq \mathcal{A}\delta(\gamma) + \bar{\mathcal{A}}\mathcal{K}_n \left(\frac{\theta_n}{4}\right)^{\omega_n - 1} \sum_{k=0}^{N-1} d_k \gamma^{t_k} + \mathcal{O}\left(\gamma^{t+N-1}\right) \quad (38)$$

where $\mathcal{O}$ denotes the order of a term in the asymptotic series and

$$d_k = \sum_{j=0}^{k} \frac{(-1)^j \sigma_n^j \left(\frac{\theta_n}{4}\right)^{k-j}}{j!(k-j)!\Gamma(k+\mu_n-j)}, \; t_k = \alpha_n(k+\mu_n)/2 - 1. \quad (39)$$

$$p_\gamma^e(\gamma) \doteq \mathcal{A}^e\delta(\gamma) + \bar{\mathcal{A}}^e \frac{a_n\sqrt{c_n}}{2} \sum_{k=0}^{N-1} d_k^e \gamma^{t_k^e} + \mathcal{O}\left(\gamma^{t+N-1}\right) \quad (40)$$

where

$$d_k^e = \sum_{j=0}^{k} \frac{(-1)^j b_n^j \left(\frac{c_n}{4}\right)^{k-j}}{j!(k-j)!\Gamma(k+2-j)}, \; t_k^e = \alpha_n(k+1)/2 - 1. \quad (41)$$

The simplified asymptotic series expressions for decoding error probability over $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channel can direly be evaluated using equations (38) and (40).

5. Numerical experiments and simulations

Though not a requirement, the performance analyses of the present model assume that the n hops of the DF-based multi-hop system are independent and identically distributed (IID) $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme variates for Case I and Case II, respectively. Furthermore, we assume that all the nodes between S and D are equally spaced. Following the standard set by [2], [37], [16] and other similar works, the maximum distance between S and D is taken as $n_{max} = 5$, upon normalizing by the distance of one hop. To further ensure fairness in comparisons, the expressions of $\bar{\gamma}$ and $\bar{\gamma}_n$ need to be multiplied by $(n/n_{max})^\beta$, where β is the propagation loss exponent which is taken as 3 for the current setup. The transmission power is assumed to be equally allocated to all the n transmitters. So, for each case respectively, we have

$$\bar{\gamma}_n = \left(\frac{n}{n_{max}}\right)^\beta \frac{e^{\mu_n\kappa_n}[\mu_n(\kappa_n+1)]^{\frac{2}{\alpha_n}} E_b/N_0}{\Gamma\left(\frac{2}{\alpha_n}+\mu_n\right)\tilde{\Phi}\left(\frac{2}{\alpha_n}+\mu_n;\mu_n;\mu_n\kappa_n\right)n}, \quad (42)$$

$$\bar{\gamma}_n^e = \left(\frac{n}{n_{max}}\right)^\beta \frac{e^{2m_n}(2m_n)^{\frac{2}{\alpha_n}-1} E_b/N_0}{\Gamma\left(\frac{2}{\alpha_n}+\mu_n\right)\tilde{\Phi}\left(\frac{2}{\alpha_n}+\mu_n;\mu_n;\mu_n\kappa_n\right)n}. \quad (43)$$

Finally, the SNR threshold for each hop is determined from the relation $\log_2(1+\gamma_{th}) = \mathcal{R}$, where $\mathcal{R}$ has been kept at 0.3 bit/s/Hz for the present work. We have validated the accuracy of the derived performance metrics via the Monte Carlo simulations. In the Monte-Carlo simulations, 10^6 random samples of $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels are generated, which are sufficient for the desired accuracy of the proposed methodology. All the computations are carried out in MATLAB® (version 2014a). The bivariate Fox H-function has not been implemented in standard packages. Nonetheless, based on linear optimization, [38] recently proposed versatile GPU-enabled MATLAB® and C/MEX codes with automated contour computation for the general multivariate Fox H-function of the type in [36] and we use a modified version of it in the present work.

Fig. 1 and Fig. 2 present the AoF vs average SNR under severe and moderate fading conditions for both Case I and Case II,

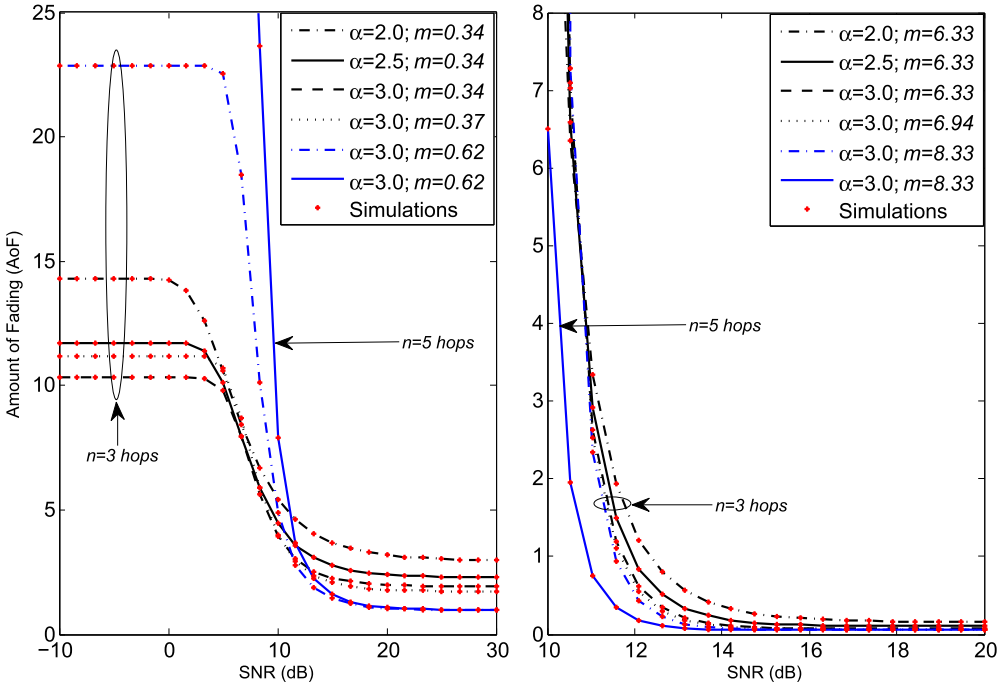


Fig. 2. AoF curves for the $\alpha-\kappa-\mu$-Extreme case showing the effects of variation in model parameter values and hop numbers under moderate and severe fading conditions similar to Case I.

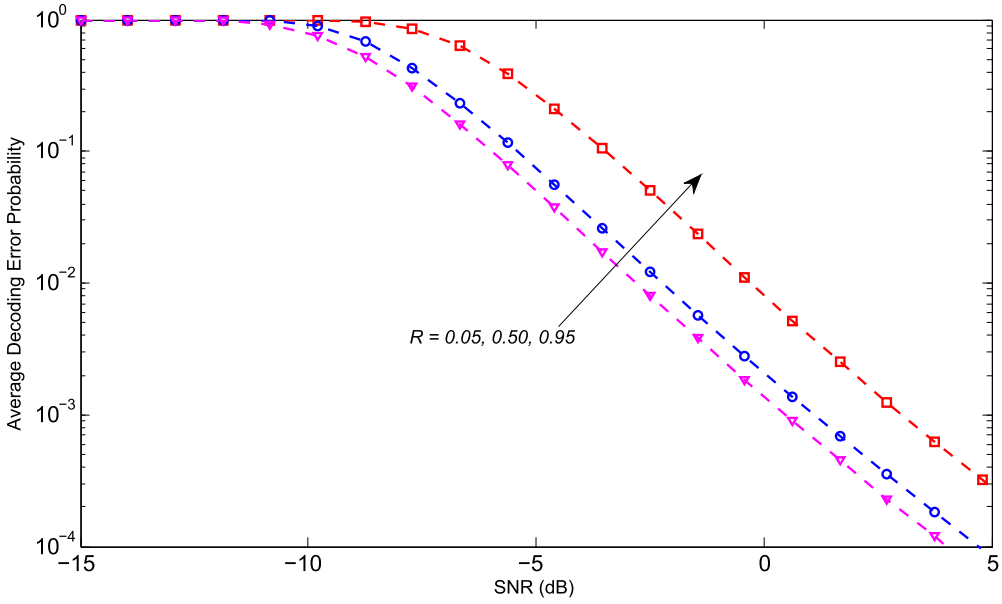


Fig. 3. Average decoding error probability for the $\alpha-\kappa-\mu$ case with $(\alpha,\kappa,\mu)=(2.5,2.6,2.25)$ and $N=200$ for a 2-hop system.

respectively. It is evident that AoF decreases with increments in values for any of the fading parameters of both models. This makes sense since the parameter m is inversely proportional to the fading severity. We also note that both models perform almost identically for moderate fading with the extreme distribution showing a slightly lower AoF values. However, the improved performance in the $\alpha-\kappa-\mu$-Extreme model over the $\alpha-\kappa-\mu$ model is clearly apparent under severe fading conditions with the latter experiencing sharper upward spikes for low SNR regimes as compared to the former. Furthermore, AoF under Case II stabilizes to a constant as SNR decreases indefinitely. While increasing the number of hops appears to lower the AoF for moderate fading conditions, for severe fading, there is a steep increase in AoF at low SNR with increases in signal power, the number of multi-clusters μ or m having little to no positive effect. Fig. 3 shows the average coding error probability for a 2-hop DF relaying system. We observe that as the transmission rate R increases, $\bar{\epsilon}$ increases, confirming the results of [33]. Fig. 4 and Fig. 5 present the channel capacity under ORA adaptive transmission for Case I and Case II, respectively. There is improvement in channel capacity for increments in any of the fading model parameter values in Fig. 4 in moderate fading. However, the improvements are very thin and even larger increments in parameter values do not cause a bigger gap between the curves. Fig. 5 depicts the ORA channel capacity of the system under extremely severe fading conditions. Here we note a considerable steeper improvement in system performance with increase

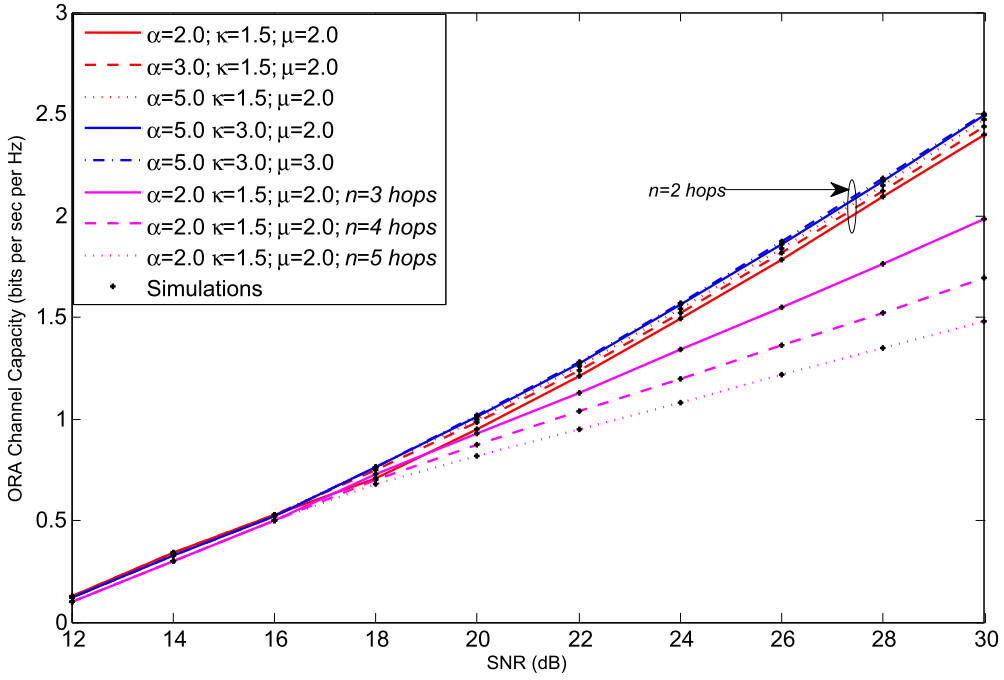


Fig. 4. $\mathcal{C}_O$ curves for the $\alpha - \kappa - \mu$ case showing the effect of increasing number relays and fading model parameter values under moderate fading conditions.

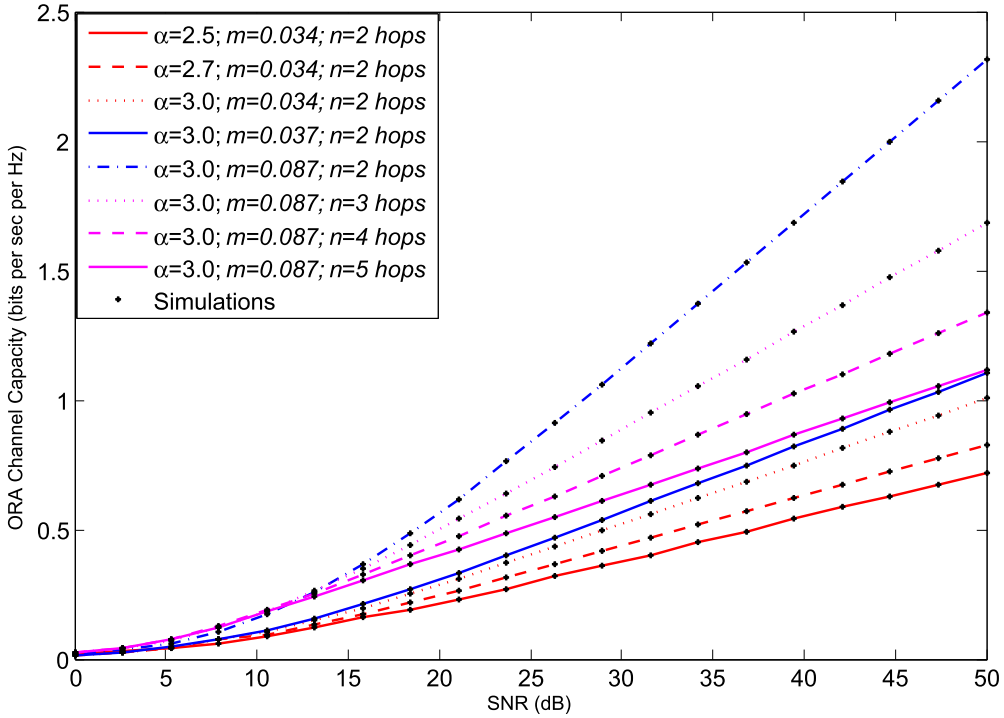


Fig. 5. $\mathcal{C}_O^e$ curves for the $\alpha - \kappa - \mu$-Extreme case showing the effects of increasing number of relays and fading model parameter values under extremely severe fading conditions.

in fading model parameter values. However, in both figures, there appears to be a diminishing reduction in channel capacity as more relays are added to the DF-based system. According to [16] this is due to the half-duplex capability of the relays which employs time-sharing multiple-access mechanism. Consequently, increasing the number of hops will require more time to deliver information from S to D, which hampers system capacity. Finally, Fig. 6 shows the channel capacity under all the adaptive transmission schemes for moderate fading conditions. It is perceived from the figure that OPRA policy has the best performance among all, while the CIFR policy has the lowest performance in the low SNR region. The TIFR policy provides better throughput than the ORA scheme in the low SNR region, but gets outperformed by ORA as the SNR value increases, as TIFR transmits only above a optimum cut-off SNR, while ORA has constant power output. The ORA power schemes converge with OPRA policy at higher SNR values, and the CIFR policy tends to converge with the TIFR policy at even higher SNR regime. The results are in agreement with the findings of [34] and prove that as expected $\mathcal{C}_P \geq \mathcal{C}_O \geq \mathcal{C}_T \geq \mathcal{C}_c$. We expect similar results to the trend set in the preceding analyses for

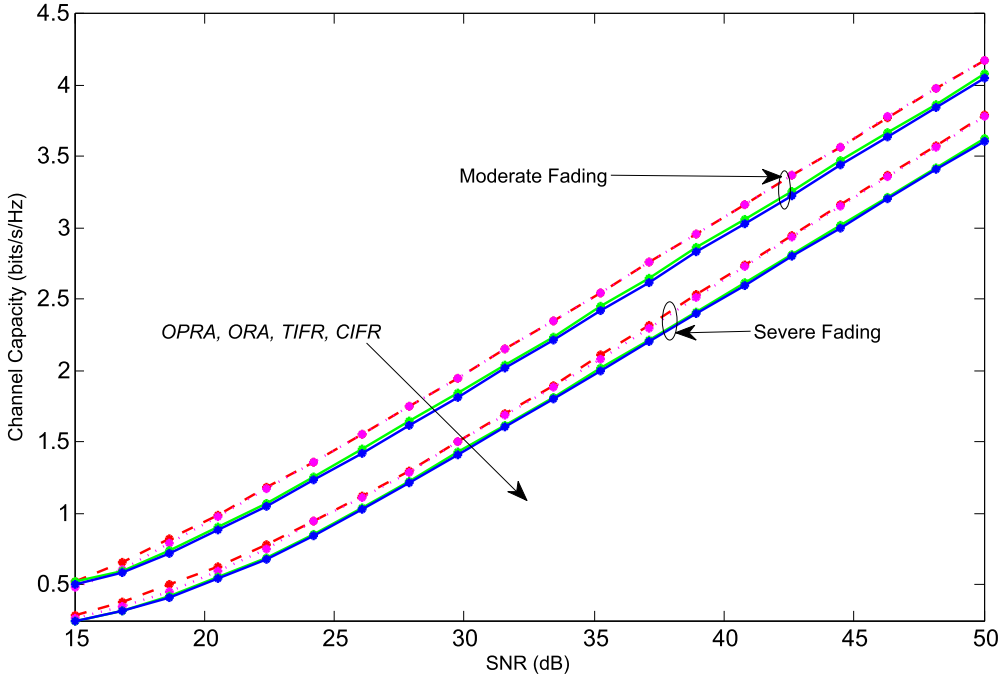


Fig. 6. Channel capacity curves for the $\alpha-\kappa-\mu$ case under various adaptive transmission schemes for a 3-hop DF-based system.

Case II and we therefore avoid displaying the figure for the sake of brevity.

6. Conclusions

In this paper, various performance measures of multi-hop system over generalized $\alpha-\kappa-\mu$ and $\alpha-\kappa-\mu$-Extreme fading channels have been examined in detail. Using the PDF-based approach, we have provided the closed-form expressions for the amount of fading, channel capacity under various adaptive schemes and decoding error probability for short packet communication for an end-to-end DF-based multi-hop relaying system. The derived expressions can be used to investigate the performance of other fading channel models by setting the associated channel parameters. Numerical results were reported to illustrate the effects of the various system and channel parameters on the communication system performance. System performance was shown to improve in a diminishing manner as the number of hop links increased under moderate fading. However, for severe fading conditions, there is little gain in the system performance. The analytical results demonstrate that a considerable improvement in system channel capacity can be realized by considering different adaptive transmission schemes. From the analyses carried out in the present work, a few points which could also be of interest for further research is to determining the number of hops at which a negligible gain in system performance kicks in so as to avoid futile additional relaying if the purpose of relaying is to improve system performance.

CRediT authorship contribution statement

It is stated that all the authors have equal contribution in the preparation of this manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Channel capacity under ORA

For a n-hop DF-based multi-hop system, the ORA channel capacity for fading channel can be calculated by averaging the capacity of an additive white Gaussian noise (AWGN), over $p_\gamma(\gamma)$, and is given as

$$\mathcal{C}_O = \frac{B}{n\ln 2}\int_0^\infty \ln(1+\gamma)p_\gamma(\gamma)d\gamma. \tag{44}$$

For Case I, we substitute (1) into (44) and set $y=\gamma^{\alpha_n/2}$ to obtain

$$\frac{n\mathcal{C}_O}{B} = \frac{\bar{A}\alpha_n\mathcal{K}_n}{\ln 2}\overbrace{\int_0^\infty \frac{\ln\left(1+y^{\frac{2}{\alpha_n}}\right)}{y^{1-\omega_n}e^{\sigma_n y}} I_{\mu_n-1}\left(\sqrt{\theta_n y}\right)dy}^{\mathcal{I}_3}. \tag{45}$$

Proceeding as before, making routine use of the definition of the Fox H-function of a single variable from [39, Eq. 1.1.1] yields the double Mellin-Barnes integral

$$\begin{aligned}\mathcal{I}_3 =& \frac{1}{(2\pi i)^2}\int_{\mathcal{L}_r}\int_{\mathcal{L}_t} \frac{\Gamma\left[\omega_n+\left(\frac{2}{\alpha_n}\right)r+t\right]\Gamma(\omega_n-1-t)}{\Gamma\left(1-\frac{\mu_n}{2}+t\right)\Gamma\left(\frac{\mu_n}{2}-t\right)} \\ &\times \frac{\pi\Gamma(1-r)\Gamma(r)\Gamma(r)\left(\frac{1}{\sigma_n}\right)^{\frac{2r}{\alpha_n}}\left(\frac{\theta_n}{4\sigma_n}\right)^t drdt}{\sigma_n^{\omega_n}\Gamma(1+r)\Gamma(\omega_n+t)\Gamma\left(\frac{\mu_n}{2}-t\right)},\end{aligned} \tag{46}$$

where $Re\,(\omega_n+2r/\alpha_n+t)>0$. Comparing (46) with [36, Eq. 2.57] results in the closed-form expression for $\mathcal{C}_O$ under Case I as shown in (24) of Table 2. Following the similar approach the ORA capacity for Case II comes out as (25) of Table 2. Using integration by parts, [35] showed that (44) has the alternative representation

$$\mathcal{C}_O = \frac{B}{n\ln 2}\int_0^\infty \frac{\left[1-P_\gamma(\gamma)\right]}{1+\gamma}d\gamma. \tag{47}$$

Therefore, substituting (2) and (5) into (47) yields expressions for channel capacity with ORA under Case I and Case II respectively as

$$\mathcal{C}_O = \frac{B\bar{\mathcal{A}}}{n\ln 2}\int_0^\infty \mathrm{Q}_{\mu_n}\left(\sqrt{2\kappa_n\mu_n}, \sqrt{2\sigma_n\gamma^{\frac{\alpha_n}{2}}}\right)/(1+\gamma)d\gamma, \tag{48}$$

$$\mathcal{C}_O^e = \frac{B\bar{\mathcal{A}}^e}{n\ln 2}\int_0^\infty \mathrm{Q}_0\left(2\sqrt{m_n}, 2\sqrt{m_n(\gamma/\bar{\gamma})^{\frac{\alpha_n}{2}}}\right)/(1+\gamma)d\gamma. \tag{49}$$

The integrals (48) and (49) can be evaluated via efficient in-build numerical integration routines of MATHEMATICA® and MATLAB®.

Appendix B. Channel capacity under OPRA

Under OPRA, the channel capacity of the end-to-end SNR γ of a n-link DF-based multi-hop system with density f_γ is defined as [35, Eq. 10]

$$\begin{aligned}\mathcal{C}_P &\triangleq \frac{B}{n}\int_{\gamma_o}^\infty \log_2\left(\frac{\gamma}{\gamma_o}\right)p_\gamma(\gamma)d\gamma,\\ &= \frac{\gamma_o B}{n\ln 2}\int_1^\infty \ln(y)p_\gamma(\gamma_o y)dy, \text{ upon setting } \gamma=\gamma_o y.\end{aligned} \tag{50}$$

The γ_o must satisfy the following relation

$$\int_{\gamma_o}^\infty (1/\gamma_o - 1/\gamma)\,p_\gamma(\gamma)d\gamma = 1. \tag{51}$$

In order to evaluate (50) for Case I, we express the PDF (1) in terms of the Fox H-function and after some simple manipulations we get

$$\begin{aligned}\mathcal{C}_P =& \mathcal{K}^*\frac{1}{(2\pi i)^2}\int_{L_r}\int_{L_t}\overbrace{\int_1^\infty \frac{\ln\left(y^{\frac{2}{\alpha_n}}\right)}{y^{1-\omega_n-r-t}}dy}^{\mathcal{H}}\\ &\times\frac{\Gamma(-r)\Gamma(\omega_n-1-t)\left(\sigma_n\gamma_o^{\frac{\alpha_n}{2}}\right)^r\left(\frac{\theta_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\right)^t drdt}{\Gamma(\omega_n+t)\Gamma\left(1-\frac{\mu_n}{2}+t\right)\Gamma\left(\frac{\mu_n}{2}-t\right)},\end{aligned} \tag{52}$$

where $\mathcal{K}^* = \left(2\gamma_o^{\frac{\alpha_n}{2}\omega_n}\pi B\mathcal{K}_n\right)/(\alpha_n n\ln 2\bar{\mathcal{A}}^{-1})$. The integral $\mathcal{H}$ can be evaluated using integration by parts as $\mathcal{H} = 2/\left[\alpha_n(\omega_n+r+t)^2\right]$ resulting in the double Mellin-Barnes contour integral

$$\begin{aligned}\mathcal{C}_P =& \frac{2\mathcal{K}^*}{\alpha_n}\frac{1}{(2\pi i)^2}\int_{L_r}\int_{L_t}\frac{\Gamma(-r)\Gamma(\omega_n+r+t)^2}{\Gamma(1+\omega_n+r+t)^2}\\ &\times\frac{\Gamma(\omega_n-1-t)\left(\sigma_n\gamma_o^{\frac{\alpha_n}{2}}\right)^r\left(\frac{\theta_n}{4}\gamma_o^{\frac{\alpha_n}{2}}\right)^t drdt}{\Gamma(\omega_n+t)\Gamma\left(1-\frac{\mu_n}{2}+t\right)\Gamma\left(\frac{\mu_n}{2}-t\right)},\end{aligned} \tag{53}$$

where $\Re(\omega_n+r+t)<0$. Finally, comparing (53) with [36, Eq. 2.57] gives a closed-form expression for the OPRA channel capacity under Case I and Case II as shown in (26) and (27) of Table 3, respectively. Applying integration by parts in (51) results in the alternative representation for the cutoff SNR relation as [35, Eq. 13]

$$\begin{aligned}&\left[1-P_\gamma(\gamma_o)\right]/\gamma_o - \int_{\gamma_o}^\infty p_\gamma(\gamma)/\gamma d\gamma = 1\\ \Longrightarrow\ &\gamma_o = \frac{1-P_\gamma(\gamma_o)}{1+\int_{\gamma_o}^\infty p_\gamma(\gamma)/\gamma d\gamma} = \xi(\gamma_o).\end{aligned} \tag{54}$$

It is clear that the solution to (54) is a fixed point of the non-linear equation $\gamma_o = \xi(\gamma_o)$. According to [35], γ_o will always lie in the interval $[0,1]$ regardless of the fading model and number of relays used for the wireless system. The explanation for the latter relies on the properties of the CDF of γ and is fully documented in Section 3 of [35]. Thus, choosing a starting point in $[0,1]$, γ_o can be computed via iterative schemes such as the Newton-Raphson method as

$$\gamma_o^i = \xi(\gamma_o^i),$$

where γ_o^i is the ith iterate of γ_o. The iterative procedure will be stopped when $|\gamma_o^{i+1}-\gamma_o^i|$ is sufficiently small. We also note again that using integration by parts on (50) yields [35, Eq. 11]

$$\mathcal{C}_P = \frac{B}{n\ln 2}\int_{\gamma_o}^\infty \left[1-P_\gamma(\gamma)\right]/\gamma d\gamma. \tag{55}$$

Therefore, we can use (2) and (5) in (55) to obtain the OPRA channel capacities for Case I and Case II respectively as

$$\mathcal{C}_P = \frac{B\bar{\mathcal{A}}}{n\ln 2}\int_{\gamma_o}^\infty \mathrm{Q}_{\mu_n}\left(\sqrt{2\kappa_n\mu_n}, \sqrt{2\sigma_n\gamma^{\frac{\alpha_n}{2}}}\right))/\gamma d\gamma, \tag{56}$$

$$\mathcal{C}_P^e = \frac{B\bar{\mathcal{A}}^e}{n\ln 2}\int_{\gamma_o}^\infty \mathrm{Q}_0\left(2\sqrt{m_n}, 2\sqrt{m_n(\gamma/\bar{\gamma})^{\frac{\alpha_n}{2}}}\right)/\gamma d\gamma. \tag{57}$$

Numerical integration routines from standard packages can be employed to evaluate (56) and (57).

References

[1] S.N. Islam, P. Sadeghi, S. Durrani, Error performance analysis of decode-and-forward and amplify-and-forward multi-way relay networks with binary phase shift keying modulation, IET Commun. 7 (15) (2013) 1605–1616.

[2] J. Cao, L.-L. Yang, Z. Zhong, Performance analysis of multihop wireless links over generalized-k fading channels, IEEE Trans. Veh. Technol. 61 (4) (2012) 1590–1598.

[3] I.B.G. Pôrto, M.D. Yacoub, On the phase statistics of the κ-μ process, IEEE Trans. Wirel. Commun. 15 (7) (2016) 4732–4744.

[4] Y. Ai, A. Mathur, L. Kong, M. Cheffena, Effective throughput analysis of α-η-κ-μ fading channels, IEEE Access 8 (2020) 57363–57371.

[5] A. Mathur, Y. Ai, M. Cheffena, G. Kaddoum, Secrecy performance of correlated α - μ fading channels, IEEE Commun. Lett. 23 (8) (2019) 1323–1327.

[6] T. Aldalgamouni, A.M. Magableh, S. Mater, O.S. Badarneh, Capacity analysis of α-η-μ channels over different adaptive transmission protocols, IET Commun. 11 (7) (2017) 1114–1122.

[7] S. Kumar, Performance of ED based spectrum sensing over α–η–μ fading channel, Wirel. Pers. Commun. 100 (4) (2018) 1845–1857.

[8] M. Bilim, Uplink communications with AWGGN over non-homogeneous fading channels, Phys. Commun. 39 (2020) 101047.

[9] M. Bilim, N. Kapucu, On the analysis of achievable rate for NOMA networks with cooperative users over $\kappa-\mu$ shadowed fading channels, Int. J. Commun. Syst. 32 (12) (2019) e4001.

[10] M. Bilim, QAM signaling over $\kappa-\mu$ shadowed fading channels, Phys. Commun. 34 (2019) 261–271.

[11] M. Bilim, A performance study on diversity receivers over $\kappa-\mu$ shadowed fading channels, AEÜ, Int. J. Electron. Commun. 112 (2019) 152934.

[12] M.D. Yacoub, The κ-μ distribution and the η-μ distribution, IEEE Antennas Propag. Mag. 49 (1) (2007) 68–81.

[13] G.S. Rabelo, U.S. Dias, M.D. Yacoub, The κ-μ extreme distribution: characterizing severe fading conditions, in: 2009 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference, (IMOC), IEEE, 2009, pp. 244–248.

[14] G. Fraidenraich, M.D. Yacoub, The α-η-μ and α-κ-μ fading distributions, in: 2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications, IEEE, 2006, pp. 16–20.
[15] P.C. Sofotasios, S. Freear, The α-κ-μ extreme distribution: characterizing nonlinear severe fading conditions, in: 2011 Australasian Telecommunication Networks and Applications Conference, (ATNAC), IEEE, 2011, pp. 1–4.
[16] F.H. Tha'er, H.B. Salameh, T. Aldalgamouni, Performance study of multi-hop communication systems with decode-and-forward relays over α- μ fading channels, IET Commun. 11 (10) (2017) 1641–1648.
[17] N. Kapucu, On the performance of dual-branch selection diversity combining in $\alpha - \kappa - \mu$ fading environments, Trans. Emerg. Telecommun. Technol. 30 (12) (2019) e3719.
[18] T.R. Rasethuntsa, S. Kumar, An integrated performance evaluation of ED-based spectrum sensing over α-κ-μ and α-κ-μ-extreme fading channels, Trans. Emerg. Telecommun. Technol. (2019).
[19] S. Kumar, M. Kaur, N.K. Singh, K. Singh, P.S. Chauhan, Energy detection based spectrum sensing for gamma shadowed α-η-μ and α-κ-μ fading channels, AEÜ, Int. J. Electron. Commun. 93 (2018) 26–31.
[20] S. Kumar, P. Chauhan, P. Raghuwanshi, M. Kaur, K. Singh, Ed performance over α-η-μ/lG and α-κ-μ/lG generalized fading channels with diversity reception and cooperative sensing: a unified approach, AEÜ, Int. J. Electron. Commun. 97 (2018) 273–279.
[21] G. Farhadi, N. Beaulieu, On the ergodic capacity of wireless relaying systems over Rayleigh fading channels, IEEE Trans. Wirel. Commun. 11 (7) (2008) 4462–4467.
[22] M.R. Bhatnagar, On the capacity of decode-and-forward relaying over Rician fading channels, IEEE Commun. Lett. 17 (6) (2013) 1100–1103.
[23] K. Ho-Van, P.C. Sofotasios, Exact bit-error-rate analysis of underlay decode-andforward multi-hop cognitive networks with estimation errors, IET Commun. 7 (18) (2013) 2122–2132.
[24] Y. Ai, M. Cheffena, On multi-hop decode-and-forward cooperative relaying for industrial wireless sensor networks, Sensors 17 (4) (2017) 1–21.
[25] D. Gunduz, M.A. Khojastepour, A. Goldsmith, H.V. Poor, Multi-hop MIMO relay networks: diversity-multiplexing trade-off analysis, IEEE Trans. Wirel. Commun. 9 (5) (2010) 1738–1747.
[26] T.R. Rasethuntsa, M. Kaur, S. Kumar, On the outage probability and BER of a DF-based multi-hop system over α-κ-μ and α-κ-μ-extreme fading channels, AEÜ, Int. J. Electron. Commun. 124 (2020) 153324.
[27] I. Gradshteyn, I. Ryzhik, Table of Integrals, Series, and Products, Elsevier Science, 2014.
[28] M.K. Simon, M.-S. Alouini, Digital Communication Over Fading Channels, vol. 95, John Wiley & Sons, 2005.
[29] Y. Polyanskyi, H.V. Poor, S. Verdú, Channel coding in the finite blocklength regime, IEEE Trans. Inf. Theory 56 (5) (2010) 2307–2359.
[30] A. Soulimani, M. Benjillali, H. Chergui, D.B. da Costa, On multihop Weibull-fading communications: performance analysis framework and applications, preprint, arXiv:1610.08535, 2016.
[31] Y. Sun, Á. Baricz, S. Zhou, On the monotonicity, log-concavity, and tight bounds of the generalized Marcum and Nuttall Q-functions, IEEE Trans. Inf. Theory 56 (3) (2010) 1166–1186.
[32] P.C. Sofotasios, K. Ho-Van, T.D. Anh, H.D. Quoc, Analytic results for efficient computation of the Nuttall-q and incomplete Toronto functions, in: 2013 International Conference on Advanced Technologies for Communications, (ATC), IEEE, 2013, pp. 420–425.
[33] P. Mary, J.-M. Gorce, A. Unsal, H.V. Poor, Finite blocklength information theory: what is the practical impact on wireless communications?, in: 2016 IEEE Globecom Workshops, (GC Wkshps), IEEE, 2016, pp. 1–6.
[34] J. Zhang, M. Matthaiou, Z. Tan, H. Wang, Performance analysis of digital communication systems over composite $\eta - \mu$/gamma fading channels, IEEE Trans. Veh. Technol. 61 (7) (2012) 3114–3124.
[35] A. Annamalai, R. Palat, J. Matyjas, Estimating ergodic capacity of cooperative analog relaying under different adaptive source transmission techniques, in: 2010 IEEE Sarnoff Symposium, IEEE, 2010, pp. 1–5.
[36] A.M. Mathai, R.K. Saxena, H.J. Haubold, The H-Function: Theory and Applications, Springer Science & Business Media, 2009.
[37] D. Dixit, P. Sahu, Performance of multihop communication systems with regenerative relays in η-μ fading channels, in: 2014 IEEE 79th Vehicular Technology Conference, (VTC Spring), IEEE, 2014, pp. 1–5.
[38] H. Chergui, M. Benjillali, M.-S. Alouini, Rician k-factor-based analysis of XLOS service probability in 5G outdoor ultra-dense networks, preprint, arXiv:1804.08101, 2018.
[39] A.M. Mathai, R.K. Saxena, The H Function with Applications in Statistics and Other Disciplines, Wiley, 1978.

Tau Raphael Rasethuntsa received his B.Sc. in Mathematics and Statistics from the National University of Lesotho, Maseru, Lesotho in 2014 and M.Sc. in Mathematical Engineering from Istanbul Technical University, Istanbul, Turkey in 2017. He is currently pursuing his PhD at McMaster University, Hamilton, Ontario, Canada. His research interests include reliability, applied statistics as well as mathematical and statistical aspects of signal processing with special focus on performance evaluation of wireless communication systems.

Manpreet Kaur received her B.Tech. in Electronics and Communication Engineering from Punjab Technical University, Punjab, India in 2005 and Master of Engineering in Electronics and Communication from Thapar University, Patiala, India in 2007. She is pursuing her PhD. from Delhi Technological University, Delhi, India. She is currently working as Member (Senior Research Staff) at Central Research Laboratory, Ghaziabad, Bharat Electronics Limited. Her research interest includes the wireless communication and cognitive radios networks.

Sandeep Kumar received his B.Tech. in electronics and communication from Kurukshetra University, India in 2004 and Master of Engineering in Electronics and Communication from Thapar University, Patiala, India in 2007. He received his PhD. from Delhi Technological University, Delhi, India in 2018. He is currently working as Member (Senior Research Staff) at Central Research Laboratory, Bharat Electronics Limited Ghaziabad, India. His research interests include the study of wireless channels, performance modeling of fading channels and cognitive radio networks. He is also serving as a reviewer for several international journals of IEEE, Springer, Elsevier etc.

Puspraj Singh Chauhan was born in UP, India. He received his Bachelor of Technology in Electronics and Communication Engineering from Uttar Pradesh Technical University, India in 2009, the M.Tech. degree in Digital Signal Processing from Govind Ballabh Pant Engineering College, Pauri, Uttarakhand, India in 2011. Currently he is working towards the PhD. degree in Electronics and Communication Engineering from Govind Ballabh Pant Engineering College, Pauri, Uttarakhand, India. His research interests include Wireless Communication and Propagation Channel Modeling.

Kuldeep Singh received his M.Tech. degree in Signal Processing from Delhi University in 2006 and Ph.D. degree in Computer Vision from Delhi Technological University, India, in 2016. He is currently an Assistant Professor with the Dept. of Electronics and Comm. Engg. Malaviya National Institute of Technology, Jaipur, India. Previously, he was a Senior Scientist with the Central Research Lab, Bharat Electronics Ltd., India. He also worked as a postdoctoral fellow at University of Alberta, Canada from Oct 2017 to Apr 2018. His research interest includes Computer Vision, Deep learning applications, Crowd Behaviour Analysis, and Biometrics.

Volume 8 Issue 2 (2020) 69-71 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Performance measurement of flexible manufacturing system in context to Bareilly camphor industry

Himani Saxena, RS Mishra, AK Madan
Department of Mechanical Production Industrial and Automobiles Engineering, Delhi Technological University Delhi-42, India.

Article Info
Article history:
Received 17 January 2020
Received in revised form
20 May 2020
Accepted 28 May 2020
Available online 15 June 2020

Keywords: Camphor manufacturing, Synthesis of material, Process flexibility

Abstract: FMS is a manufacturing system in which there is flexibility that allows the system to react to changes, whether predicted or unpredicted. It is a highly integrated manufacturing system. Oriental Aromatics Ltd. (Formerly Camphor & Allied Prod. Ltd.) is a completely integrated flavour, fragrance and aroma chemical manufacturing companies in the world today. Its three manufacturing centers are spread across the west and north of India, in Ambernath in Maharashtra, Baroda in Gujarat and Bareilly in Uttar Pradesh. It manufactures synthetic camphor, which can be utilised in several different ways.
Camphor has several pharmaceutical uses. While it is most commonly used as an anti-inflammatory and decongestant, it also has several other therapeutic and healing effects. The World Health Organization's (WHO) Good Manufacturing Practice (GMP) is a system for ensuring that products are consistently produced and controlled according to quality standards. It is designed to minimize the risks involved in any pharmaceutical production that cannot be eliminated through testing the final product. In this paper, a concept and implementation of the measuring and analysis of performance measures of FMS is applied.

1. Introduction

Camphor is one of the products used during religious prayers. This is a part of worship since ancient days. These are available in different size and shape. Camphor Tablets is a mass consumption item and is practically being consumed in all religious festivals in temples.

The consumption is directly proportional to the persons visiting any religious place. Hence, consumption will increase with the increase in faith in "GOD" and propose concern can sell its product very easily. In addition to the religious demand, it is also required in Ayurveda medicines. The overall properties of camphor differ to a large extent from thermoplastics and other materials. Therefore, it is important to consider properties while designing the mould which will affect mould parameters.

2. Manufacturing Process

The brief description of the manufacturing process of camphor tablet is given below:

a) Firstly preparation of bornyl chloride by the action of dry hydrogen chloride on α-pinene.
b) Isomerization of α-pinene to camphene.
c) Camphor synthesis is hydrolysis of isobornyl acetete with aqueous solution of sodium hydroxide. Then resulting isoborneol subjected to dehydrogenation under catalysts and target camphor is obtained.
d) Camphor powder is given heat treatment to remove the dampness available in the powder.
e) Treated powder is filled into the hopper of the tablet machine.
f) Tablet machine produce the tablet of specific shape and weight. The shape & weight of tablet will be according to the dies.
g) Five tablets are packed in one wax wrapper.
h) These box are packed in the cartoon boxes and dispatched.

Quality of the product must be as per customers demand and according to BIS standards.

Table 1: Conversion of α-pinene and selectivity to camphene with TCA/Neutral zeolite catalyst

Catalysts	Reaction time (min)	320 W X (%)	320 W S (%)	480 W X (%)	480 W S (%)	640 W X (%)	640 W S (%)	800 W X (%)	800 W S (%)
TCA/ H-Natural zeolite	15	3,47	33,97	12,26	27,73	23,25	28,26	29,83	22,90
	30	3,11	41,73	18,13	26,86	35,60	27,78	70,85	23,60
	45	5,40	33,88	29,75	26,35	66,88	26,30	88,48	19,43
	60	8,68	26,27	42,37	26,22	84,97	24,88	97,05	13,73
	90	13,46	22,43	76,01	24,26	98,99	16,90	98,85	6,59

X =Conversion of α-pinene, S = selectivity to camphene

***Corresponding Author,**
E-mail address: mechhimani@gmail.com
All rights reserved: http://www.ijari.org

Table 1 presents α-pinene conversion values (%X) and selectivities to camphene for the different treatments performed. When the TCA/Natural zeolite is exposed to a more severe power of microwave treatment, 320, 480, 640, 800 W the conversion increases but selectivities to camphene is decreases. The main product were camphene, α-terpinene, limonene, p-cimene, γ-terpinene, and terpinolene.

3. Working standards of the Camphor industry.

Table 2: Area used for FMS camphor industry in Bareilly.

1.	Covered area	Sq. Ft.	500
2.	Uncovered area	Sq. Ft.	500
3.	Total area	Sq. Ft.	1000
4.	Whether constructed or Rented		Rented
5.	If constructed, constructed Value	Rs	N.A.
6.	If Rented, Rental value (per month)	Rs	2000

Table 3: Types of Machines and equipments used.

S.N.	Description	Qty.	Value (Rs.)
1.	S.S.Mixer	1	125000
2.	Multi Punch Tablet Machine	1	
3.	Dies	L.S.	
4.	Wax wrapping machine	1	
5.	Weighing Balance	1	
6.	Hand Tools	L.S.	
7.	Sales Tax, Freight & Insurance etc.		12500
	Total		137500

Table 4: Raw material

S.N.	Particulars	Quantity (Kg)	Value (Rs)
1.	Camphor Powder	1000	100000

2.	Packaging Material		2000
		Total	102000

Table 5: The staff and Labour per month

S.N.	Particulars	Qty	Rate	Value (Rs)
A	Administrative And Supervisory			
(i)	Manager	1	3000	3000
(ii)	Peon/ Chowkidar	1	2000	2000
B	Technical (Skilled-Unskilled)			
(i)	Skilled Worker	1	3000	3000
(ii)	Unskilled Worker	2	2000	4000
	Sub-Total			12000
	Plus perquisites @ 30% of salaries			3600
	TOTAL			15600

Table 2, 3, 4 and 5 presents the working standards of the camphor industry visited for the reference of FMS study. Table 2 represents area used for FMS camphor industry in Bareilly and table 3 represents types of Machines and equipment's used. Table 4 reprrsents Raw material used in the camphor industry of Bareilly and table 5 represents staff and Labour per month working in the factory and the salaries.

4. Controlling Parameters

Compression molding method is used to develop superior and desired properties of the composite. All the three dimensions of the model (pressure, temperature and time of application) are critical and have to be optimized effectively to achieve tailored product as every dimension of the model is equally important to other one. If time of application of these factors (pressure and temperature) is not sufficient (high or low), it may cause any of defects associated with insufficient pressure or temperature.

The other manufacturing factors such as mold wall heating, closing rate of two matched plates of the plates and de- molding time also affect the production process. The overall properties of camphor differ to a large extent from thermoplastics and other materials.

Therefore, it is important to consider properties while designing the mould which will affect mould parameters. Dimensional accuracy of molded camphor parts is very important and in such cases, the design of the mould must allow for shrinkage of the parts. For analyzing the above properties of given camphor sample and mould parameters, camphor compression mould is specifically designed. Hydraulic system should be avoided because hydraulic oil leakage may affect the production vigorously. Hydraulic oil may react with camphor and it is highly undesirable.

As the camphor is chemically reactive towards the hydraulic oil thus we cannot use hydraulic press for the operation which is generally prefer for press operation on the other hand we have option of pneumatic press but as the camphor is sublimatory substance which will vaporize the camphor material when come in contact with air and pneumatic system also have high initial investment as well as complicated components like pump ,direction control valve, actuators etc. and those are need to be handled with skilled and experience operator so the press operation must be carried out using the mechanical system which includes the various mechanisms like cam and follower, speed reduction using gears, reciprocating motion for press operation uses the slider crank mechanism.

4.1 Working Capital (For One Month)

Table 6: Working capital of industry.

Sl.No.	Description	Amount(Rs)
1	Raw material	102000
2	Salaries & Wages	15600
3	Other Expenses	11500
	Total	129100

4.2 Total Capital Investment

Table 7: Total capital investment

Building & Other Civil Works	-
Machinery & Equipment	137500
Working capital for one month	129100
Total	266600

4.3 Cost of Production (Per Annum)

Table 8: Cost of production (per annum)

Total recurring cost per year	1549200
Depreciation on machinery & Equipment	13000
Interest on total investment @ 10%	27000
Total	1589200

4.4 Sales Proceeds(Per Annum)

Table 9: Sales proceeds per annum

S. N.	Item	Quantity(Kg)	Value (Rs.)
1.	Camphor Tablets	12000	1800000
	Total		1800000

4.5 Profitability

Table 10: Profitability (before income tax).

1.	Annual Gross Profit	210800
2.	% of Profit on Sales	11.71%
3.	Break Even Analysis	
3.1	Annual Fixed Cost	325200
3.2	Annual Sales	1800000
3.3	Annual Variable Cost	1224000
3.4	Break Even Point	56.46%

Table 6.7.8.9 and 10 represents the controlling parameters taken by industry to optimally to achieve the efficient product.

4.5.1 Break-Even Analysis

(% of Total Production envisaged)

Annual fixed cost X 100

---------- ---------- ---------- = %

Annual sales – Annual variable costs

4.5.2 Production Targets

The final product of camphor are of two types.

Basis of estimation: Single Shift basis (6 hours per shift)
Type 1 Camphor: 96% purity
Type 2 Camphor: 90% purity
According to the desired properties of product, mechanical drive system is preferable for machine.

5. Assumption for Generating Project Profitably

Table 11: Assumptions for generating project.

1	Number of Working Days in a year	300 Days
2	Number of Shifts in a day	1 One
3	Hours in a Shift	8 hours
4	Plant Capacity	Consider on Average production capacities of plant.
5	Raw material Estimates	Based upon product Mix
6	Depreciation	Straight Line Method
7	Manpower	According to project Requirement
8	Rent estimate	On the basis of current market price of the area.
9	Potential Area of Marketing the products	Households & related industries of the area.
10	If project is funded, term loan would be	60-80% of Total investment
11	Moratorium Period	6- 12 months
12	Repayment Period	5-7 years

Table 11 represents the further assumptions that could generate more profit to the industry concerning with working hours and output generated with respect to time.

6. Conclusions

Using various design software we can conclude that punch for the desired tablet product is withstanding in all the working conditions. In this various moulding techniques used for manufacturing of polymer products are studied. Concept and implementation of the measuring and analysis of performance measures of FMS is satisfied in Bareilly camphor industry.

Acknowledgment

Authors thank to Mr. Satish Ray, Director Operations of Camphor plant Bareilly, Mr. Alam, Engineer at Camphor plant Bareilly for the support and my Guides, friends, family also the almighty.

References

[1]. OP Khanna, M Lal. A text book of Production Technology, Dhanpat Rai Publications (P) Ltd., New Delhi, 2001.
[2]. Camphor manufacturing unit library, Bareilly camphor industry.
[3]. JG Gurvich. On the theory of heterogeneous catalyst. J. Russ. Phys. Chem. Soc. 48 (1), 1915, 837-856.
[4]. VE Tishchenko, GA Rudakov. Catalytical method of preparation of camphene from turpentine. Russ. J. Appl. Chem. 6 (1), 1933, 691-70.
[5]. NS Patel. Modeling Design and Analysis of Cam and Follower, International Journal of Engineering Studies and Technical Approach.
[6]. PP Thakare, VL Kadlag. Design and Analysis of Helical Compression Spring for Special Purpose Application, IJARIIE.
[7]. MM Billah, MS Hossain, MR Islam, MA Rahman. Selection of V-Belt for Power Transmission in Machinery.
[8]. PS Valsange. Design of Helical Coil Compression Spring, International Journal of Engineering Research and Applications.
[9]. DM Stiel, CR Browski, E Neugroschel, JJ Williams. Tablet Press US Patent 4,100,598, 1978.
[10]. L Norton Robert. Design of machinery, Tata McGraw-Hill Publishing Company Ltd., New Delhi, 2005.
[11]. SM Bruno, M Pillinger, FE Kuhn, IIS Goncalves, AA Valente. Isomerization of αpinene oxide in the presence of methyltrioxorhenium (VII). Catalysis Communications. 35, 2013, 40- 44.
[12]. N Wijayati, D Pranowo, P Triyono. The Acid Catalyzed Reaction Of α -Pinene Over Y-Zeolite. Indonesian Journal of Chemistry, 13(1), 2013, 59-65.
[13]. NT Wijayati. Handayani, Supartono. Isomerization of α-pinene using Zirconiua/Natural Zeolite Catalysts. Asian Journal of Chemistry. 29 (8), 2017, 1705-1708.
[14]. NA Comelli, N PonziE, MI dan Ponzi. α-Pinene Isomerization to Camphene Effect of Thermal Treatment on Sulfated Zirconia. Chemical Engineering Journal, 117(2), 2006, 93-99.
[15]. J Wang, W Hua, Y Yue, Z Gao. MSU-S Mesoporous Material: An Efficient Catalys For Isomerization of α-Pinene. Bioresource Technology,101, 2010, 7224-7230.

INTERNATIONAL JOURNAL OF GREEN ENERGY
https://doi.org/10.1080/15435075.2020.1847115

Taylor & Francis
Taylor & Francis Group

Check for updates

Performance evaluation of solar based combined pre-compression supercritical CO2 cycle and organic Rankine cycle

Yunis Khan and Radhey Shyam Mishra

Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

ABSTRACT

The present study deals with performance analysis of solar power tower powered combined pre-compression supercritical CO_2 cycle and organic Rankine cycle (ORC) for recovering waste heat. Results showed that net power output and thermal efficiency of pre-compression cycle was improved by 4.51 and 4.52%, respectively, using ORC. Highest thermal, exergy efficiency, and power output of the combined cycle increased with solar irradiation having highest values 51.83, 74.06%, and 278.5 kW at 1000 W/m^2 of solar irradiation using R227ea. Waste heat recovery ratio increased with heat exchanger effectiveness. Its highest value was found as 0.5673 at 0.95 of effectiveness based on R227ea.

ARTICLE HISTORY
Received 19 July 2020
Accepted 9 October 2020

KEYWORDS
Thermodynamic analysis; pre-compression sCO_2 cycle; waste heat recovery; solar power tower; organic Rankine cycle

1. Introduction

World's demand for energy is rising daily basis due to industrial expansion and growth in population (Guo et al. 2016; Yagli 2020). Consequently, fossil fuel consumption is continuously raising leading to increased carbon emissions. However, fossil-fuel stocks are steadily decreasing. This leads to the challenge of finding clean and reliable energy resources (Maali and Khir 2020; Wei et al. 2014). There are currently numerous renewable energy resources being used to generate clean and environmentally friendly power such as, solar, wind, biomass, and geothermal energy. Because of its availability, low cost, and noise-free operation, solar energy is more suitable in power generation among these renewable energy resources (Kabira et al. 2018; Sarver, Al-Quraghuli, and Kazmerski 2013).

Solar collectors are used to collect sunlight for cooling, heating, and power generation (Al-Sulaiman 2014; Bishoyi and Sudhakar 2017; Zhai et al. 2009). Tian and Zhao (2013) reviewed solar collectors and founded that flat plate collectors are suitable for application at low temperatures up to 100°C. However, concentrated solar collectors like parabolic trough collectors, dish collectors, heliostats, etc. are used for high-temperature applications. Kalogirou (2004) recorded parabolic trough collector (60–500°C), dish collector (750–1000°C), and heliostat (150–1500°C) working temperature range. Among other solar technologies, Concentrated Solar Power (CSP) technology plays an key role in solving the current and future question of power generation in tropical countries through the use of solar heat that is readily accessible in nature (Fuqianga et al. 2017).

Solar power tower (SPT) is the latest technology amongst the various CSP technologies. SPT system consists of number of complex sub-systems such as receiver, 75–150 m high tower, thermal storage system (optional), heliostat field with per heliostat area of 50–150 m^2 and power conversion system. Solar radiation is centered on receiver by heliostat field where it is used to produce high-temperature heat for power generation through a high energy cycle or for industrial process supply (Barlev, Vidu, and Stroeve 2011; Farges, Bézian, and El-Hafi 2018).Several studies were performed on the SPT- driven cycles such as combined recompression sCO_2 cycle and transcritical CO_2 cycle (Khatoon and Kim 2020), sCO_2 Brayton cycle (Alsagri, Chiasson, and Gadalla 2018), recompression sCO_2 with and without main compressor inter cooling (Maa et al. 2019), triple-combined cycle (Shukla et al. 2018), multi generation combined cycle (Abid, Adebayo, and Atikol 2019).

Supercritical CO_2 (sCO_2) is the state of CO_2 above its critical values (7.38 MPa, 30.98°C) Further, sCO_2 cycle is the cycle that can utilize the heat from various heat sources such as solar thermal energy, geothermal, coal power and natural gas (Ahn et al. 2015). Few researches were published articles on the basis of sCO_2 such as Khan and Mishra (2020) conducted a study on the solar parabolic trough collector integrated combined partial heating sCO_2 cycle and ORC. They also compared the performance of the combined system with and without the solar. As a result, the application of the ORC to the existing partial heating cycle enhanced the thermal efficiency by 4.47%. Further, they investigated the effect of solar radiation on performance of the combined system. They concluded that with the solar irradiation, the system's thermal and exergy efficiency increased. Finally, they also observed that R1233zd (E) is the best performing fluid among the working fluids considered and parabolic trough collector alone responsible for the 62.93% of overall cycle's exergy destruction rate. Besarati and Goswami (2014) performed a comparative analysis of the various configurations of the sCO_2 cycle such as recompression, simple recuperated, and partial cooling cycle. They first compared the systems standalone and used ORC as the bottom cycle thereafter to recover the waste heat from all three cycles and compared it again. They concluded that the combined cycle achieved more than 50% thermal efficiency as

CONTACT Yunis Khan yuniskhan21@gmail.com Department of Mechanical Engineering, Delhi Technological University, New Delhi 110042, India.
© 2020 Taylor & Francis Group, LLC

well as the recompression cycle being the best performing cycle among the cycles considered. Kim et al. (2016a) also performed a comparative analysis on 12 different sCO_2 cycle configurations such as simple recuperated cycle, single-heated cascade cycle, recompression, partial heating, pre-compression, dual cascade cycle, single-heated cascade cycle with intercooler, intercooled dual-heated cascade cycle, dual heated and flow split cycles for triple heating cycle, dual expansion cycle and partial recuperation cycle as landfilling gas turbine bottoming. They found various results one the conclusion obtained from their results was that pre-compression and recompression cycle had high thermal efficiency than other considered cycles. They also suggested for future research on combination of both two cycles. Yu et al. (2020) considered four configurations of the sCO_2 cycles such as pre-compression, simple recuperated, recompression cycle, and split flow recompression cycle, for recovering the waste heat from the internal combustion exhaust. Further, exhaust heat recovery ratio and thermal efficiency of the four systems were calculated. It was found that the maximum exhaust heat recovery ratio for pre-compression and recuperation of the sCO_2 cycle were achieved at 5.8 and 7.65 MPa, respectively. Furthermore, split flow cycle was found as best performing cycle with highest recovery ratio was 24.75%. Neises and Turchi (2019) examined the cost, design, and performance of the three sCO_2 cycle configurations, such as recompression cycle, simple cycle and partial cooling cycle, driven by the molten salt solar power tower. As a result, highest thermal efficiency was achieved in case of the recompression cycle followed by partial cooling and a simple cycle. Partial cooling was found to be more costly than other cycles due to the high demand for turbo machinery. Yang, Yang, and Duan (2020) analyzed the performance of the four sCO_2 Brayton cycles including reheating cycle, simple recuperative cycles, intercooling cycle and recompression cycle under the part load conditions. They concluded that the reheating cycle was more efficient than the simple recuperative cycle also under part load condition intercooling cycle and recompression cycle performed better than the reheating cycle. But intercooling cycle was only more efficient than recompression cycle when the load exceeded by 60%.

Furthermore about the pre-compression cycle, the pre-compressor is used between the main compressor and turbine in a simple Brayton cycle to make the outlet pressure of turbine independent of the compressor inlet pressure. Pre-compression layout requires one more heat exchanger to avoid pinch problems. When the temperature difference is reduced to a minimum in the recuperator, additional compression is used for further regeneration (Kulhanek and Dostal 2011). Benefit to this layout is addition of the one more compressor that makes place for further regeneration (Al-Sulaiman and Atif 2015). Additional compression increases the fluid temperature, also the specific heat of the fluid increases due to high compression. Thus, regeneration process can be continued and more heat is available for regeneration as compared to the simple Brayton cycle. This extra heat therefore leads to an increase in the average heat-addition temperature, while reducing the temperature of heat-rejection. This results in increase in efficiency over the basic Brayton cycle (Dostal 2004). One more reason to choose the pre-compression cycle over the other cycles is that it has same number of components with recompression cycle and does not have flow split. Also highest thermal efficiency can be obtained by reducing the waste heat through the two recuperations among the other sCO_2 Brayton cycles (Kim et al. 2016a). If Pre-compression cycle carefully optimized, its efficiency can be improved up to 6% over the Brayton cycle. (Angelino1968).

Apart from this, ORC is the latest technology for the recovery of low temperature waste heat. ORC is used as a bottoming cycle in the current scenario. There are a few studies to support this comment such as Shaaban (2016) carried out a solar-integrated combined cycle study in which ORC and steam Rankine cycle were used as a bottoming cycle for the recovery of waste heat. The author considered 15 working fluids in the ORC and found that R1234ze(Z) was the best fluid in terms of thermo-economic, environmental, and safety considerations. Hoang (2018) reviewed for application of the ORC as bottoming cycle to the exhaust waste heat recovery from diesel engines. They found that ORC can achieve up to 25% thermal efficiency, while combined system with diesel engine achieved up to 90% thermal efficiency. Song et al. (2018) conducted a performance and optimization analysis of the combined cycle of sCO_2 and ORC. The effects of topping cycle recuperative ratio, pressure ratio on the performance of the bottom were examined. They resulted in an improvement in the thermal efficiency of the existing sCO_2 cycle by adding the ORC system. ORC system recovered completely residual heat. Singh and Mishra (2018a) investigated the thermal performance of the combined sCO_2 cycle and the ORC powered by the parabolic trough solar collector. ORC was used as a bottoming cycle to recover the waste from the topping cycle. They concluded that R407C-based combined cycle yielded a maximum thermal and exergy efficiency of 43.49 and 78.07% at 0.95 kW/m^2 of solar irradiation, respectively. Li et al. (2018) performed an optimization analysis as well as the comparison of single and dual pressure evaporation ORCs. They found that the work output by dual pressure evaporation ORC was more than the single pressure evaporation ORC. Furthermore, as heat source temperature decreased maximum work output increased, and the maximum increments are 21.4–26.7% for nine working fluids.

From the literature survey, it is observed that performance of the pre-compression sCO_2 cycle can be enhanced further by incorporating a low temperature bottoming organic Rankine cycle. However, detailed thermodynamic analysis of the combined pre-compression sCO_2 cycle and ORC is not present in the literature. Therefore, present study deals with thermodynamic analysis of the combined power generation cycle driven by solar power tower. Exergy, thermal efficiency, net output power, and waste heat recovery ratio were considered as performance parameters of the proposed system. A computer program in engineering equation solver software was made to simulate the system. The first thermal performance of the combined cycle was compared with the standalone pre-compression cycle. Then, parametric analysis of the proposed combined cycle was conducted. The effects of the system variable such as solar irradiation, inlet temperature of main turbine, inlet pressure, and temperature of main compressor, inlet pressure of pre-compressor, heat exchanger 2 effectiveness and

low-temperature recuperator effectiveness on the system performance were investigated. Analysis and design of solar power tower is out of scope of current study.

2. Description of proposed system

The current model of combined cycle consists of two thermodynamic cycles such as pre-compression sCO_2 cycle as toping cycle and ORC as bottoming cycle for recovering waste heat from the topping cycle. Schematic diagram of the current combined system powered by the solar power tower is shown in Figure 1. Heat transfer fluid (HTF) (molten salt) circulates in the SPT field. This HTF supplies heat to the topping cycle through the heat exchanger 1 (HX1) (process a-b). Pre-compression cycle takes heat from the SPT through the heat exchanger 1 (process 9–1). Heated sCO_2 which is circulating in pre-compression cycle, expanded in the main turbine (MT) to get power out (Process 1–2). Expanded sCO_2 have much amount of heat which is recuperated in the high-temperature recuperator (HTR) to heat sCO_2 stream incoming to the HX1 (Process 2–3). After the HTR sCO_2 stream temperature becomes lower and HTR streams temperature difference reduced. To avoid pinch point problem sCO_2 stream is pre-compressed in the pre-compressor (PC) (Process 3–4). After the PC temperature and pressure increased, then sCO_2 stream goes to the low-temperature recuperator (LTR) where heat is recuperated by the low-temperature stream going to HTR (Process 4–5). Still sCO_2 stream have much amount of heat, goes to the heat exchanger 2 (HX2) where this amount of heat have given to the ORC through HX2 (process 5–6). Again sCO_2 compressed in MC to get inlet pressure of MT (process 6–7) after the MC sCO_2 stream goes to the LTR (process 7–8). After that sCO_2 stream goes to the HTR to regenerate (8–9). Finally, sCO_2 stream again goes to the HX1 to complete the cycle. Now come across to the ORC. ORC working fluids take heat through the HX2 by the exhaust of topping cycle and then expanded in organic turbine (OT) (process 10–11). Then fluid stream goes to the condenser to reject the heat (process 11–12). Fluid stream pressure again increased through the pump (process 12–13). Finally, again it goes to the HX2 to get the heat and cycle completed. Temperature–entropy (T-s) diagram of the proposed model is given in the Figure 2 for better understanding of the model.

3. Thermodynamic analysis

3.1. Assumptions

Performance analysis of the SPT powered combined cycle was performed considering following assumptions to support the simulation; (1) All system components are in equilibrium and steady state conditions. (2) Pressure and friction loss are neglected in each component. (3) All thermodynamic processes are polytropic. (4) Energy due to height and velocity of the each component is neglected (5) Heliostat and receiver parameters have kept constant and assumed input data to support the simulation are given in Table 1. (6) Molten salt temperature inlet to the HX1 has been taken 700°C (Reyes-Belmonte, Sebastián, and Romero 2016). (7) Due to thermal losses, inlet temperature of MT is 50 °C less than molten salt temperature inlet to the HX1.

3.2. Thermal modeling of SPT

Thermal modeling equations of the proposed system were developed in this part based on the conservation of exergy and energy equations, taking into consideration of assumptions those are made in above section. Also each component has been treated as control volume.

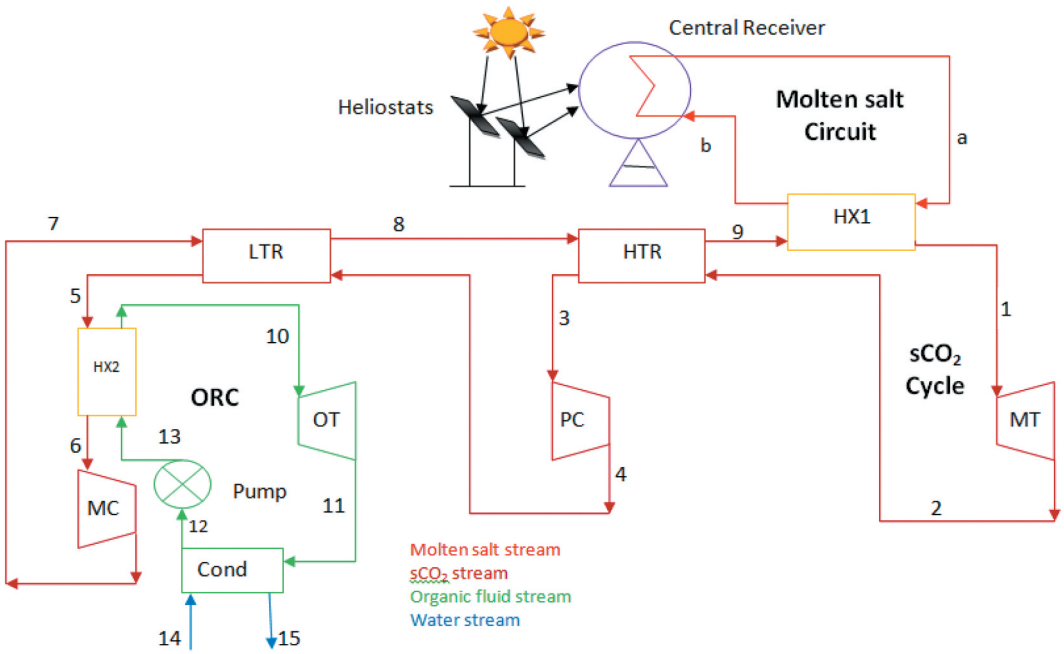


Figure 1. Schematic diagram of solar based combined cycle.

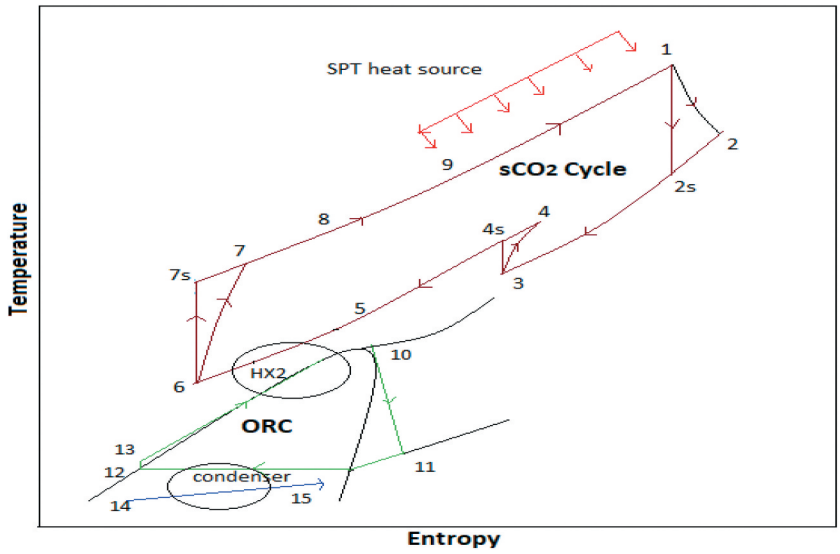


Figure 2. Temperature – entropy (T-s) diagram of proposed system.

Table 1. Input parameters for simulation of the proposed model.

Geometric and operating parameters for SPT	
Direct normal irradiation	970 W/m^2 (Khatoon and Kim 2020)
Sun temperature	4500 K (Bejan, Kearney, and Kreith 1981)
Solar's multiple	2.8 (Khatoon and Kim 2020)
Efficiency of heliostat	58.71% (Ho and Iverson 2014)
Number of heliostat	141 (Wang et al. 2018)
Heliostat's total mirror area	9.04 × 7.89 (Khatoon and Kim 2020)
Initial temperature difference	15 K (Khatoon and Kim 2020)
Solar receiver's temperature approach	423.15 K (Ho and Iverson 2014)
Concentration ratio	900 (Ho and Iverson 2014)
Convective heat loss coefficient	10 W/m^2-K (Ho and Iverson 2014)
Tower height	74.62 m (Wang et al. 2018)
Convective heat loss factor	1 (Ho and Iverson 2014)
View factor	0.8 (Ho and Iverson 2014)
Absorptance	0.95 (Ho and Iverson 2014)
Thermal emittance	0.85(Ho and Iverson 2014)
Input parameters for combined cycle	
Main turbine inlet pressure	25 MPa (Neises and Turchi 2014)
Main turbine inlet temperature	650 °C (Neises and Turchi 2014; Khatoon and Kim 2020)
Pre-compressor inlet pressure	5.6–6.8 MPa (Reyes-Belmonte, Sebastián, and Romero 2016; Neises and Turchi 2014)
Main compressor inlet pressure	7–10.5 MPa (Reyes-Belmonte, Sebastián, and Romero 2016; Neises and Turchi 2014)
Main compressor inlet temperature	32–38 °C (Reyes-Belmonte, Sebastián, and Romero 2016; Neises and Turchi 2014)
Main turbine isentropic efficiency	0.88 (Neises and Turchi 2014; Reyes-Belmonte, Sebastián, and Romero 2016)
Main compressor isentropic efficiency	0.85(Neises and Turchi 2014; Reyes-Belmonte, Sebastián, and Romero 2016)
Pre-compressor isentropic efficiency	0.85(Neises and Turchi 2014;Reyes-Belmonte, Sebastián, and Romero 2016)
Heat exchanger effectiveness	0.95(Besarati and Goswami 2014)
HTR and LTR effectiveness	0.95 (Besarati and Goswami 2014)
sCO_2 topping cycle mass flow rate	1.5 kg/s
Mass flow rate in bottoming ORC	2.5 kg/s
Inlet pressure ORC turbine	3 MPa (Khan and Mishra 2020)
ORC turbine's Isentropic efficiency	0.8 (Song et al. 2018)
ORC pump's Isentropic efficiency	0.7 (Song et al. 2018)

Direct solar heat incidence upon heliostat field is defined as (Khatoon and Kim 2020; Shukla et al. 2018)

$$\dot{Q}_{solar} = G_b \cdot A_h \cdot N_h \quad (1)$$

Where, G_b is the solar irradiation per unit area also termed as direct normal irradiation (DNI), $A_h$is single heliostat area (m^2) and N_h is the heliostats number. However, due to heliostat efficiency, some of that heat is lost in the surroundings. The amount of actual heat obtained through the heliostat field is therefore specified as (Khatoon and Kim 2020; Shukla et al. 2018);

$$\dot{Q}_h = \dot{Q}_{solar} \cdot \eta_h \quad (2)$$

Where, η_h is the efficiency of heliostat. This amount of heat is directed to the solar receiver where the heat transfer fluid flows. But in the atmosphere a part of the heat is lost. The heat available at the solar center receiver is therefore determined as (Khatoon and Kim 2020; Shukla et al. 2018);

$$\dot{Q}_r = \dot{Q}_h \cdot \eta_r = \dot{Q}_h - \dot{Q}_{loss,r} \quad (3)$$

Where, η_r is the receiver thermal efficiency, is defined as (Khatoon and Kim 2020);

$$\eta_r = \alpha - \frac{\zeta \cdot f_{view} \cdot \sigma \cdot T_R{}^4 + h_{conv} \cdot f_{conv} \cdot (T_R - T_{air})}{G_b \cdot \eta_h \cdot CR} \quad (4)$$

Where, T_R is the solar receiver surface temperature and CR is concentrated ratio. ζ is the solar emittance. To calculate heat loss, this can be approximated as (Khatoon and Kim 2020);

$$T_R = T_1 + +\delta T_R \quad (5)$$

Where, T_1 is the turbine's inlet temperature and δT_R is approach temperature of solar receiver.

The operating and geometric parameters of the solar receiver and the heliostat field are listed in Table 1.

Furthermore, exergy of any system can be explained as maximum work obtainable from the system when system is brought to its dead conditions. Control volume exergy balance equation can be determined as (Cengel and Boles 2004);

$$\sum \left(1 - \frac{T_0}{T_Q}\right) \dot{Q}_j - \dot{W}_{c\cdot v} - \sum (\dot{m}_i E_i) - \sum (\dot{m}_e E_e) - \dot{ED} \cdot = 0 \quad (6)$$

Where, $\dot{ED}$ is the exergy destruction rate and subscript j refers to thermal property at particular state. Solar exergy inlet to the combined system is determined as (Khatoon and Kim 2020);

$$\dot{E}_{solar} = \left(\frac{\dot{Q}_r}{\eta_h \cdot \eta_r}\right) \cdot E_s \quad (7)$$

Where, E_s is the dimensionless maximum useful work obtained from the solar irradiation. E_s is expressed as (Khatoon and Kim 2020);

$$E_s = 1 + \frac{1}{3}\left(\frac{T_0}{T_{su}}\right)^4 - \frac{4}{3}\left(\frac{T_0}{T_{su}}\right)(1 - \cos\beta)^{1/4} \quad (8)$$

Where, T_{su} and T_0 are the sun and reference temperature, respectively. β is the sun's disc subtended half cone angle. Its value has been taken 0.005 rad on solar energy limiting efficiency (Parrott 1978). Further, in the receiver, useful exergy obtained by the molten salt is defined as

$$\dot{E}_r = \dot{m}_{ms} \cdot Cp_{ms} \cdot \left[(T_b - T_a) - \left(T_0 \cdot \ln\frac{T_b}{T_a}\right)\right] \quad (9)$$

Further chemical exergy of the system is constant throughout. After neglecting energy due to velocity and height, specific physical exergy at j^{th} point is defined as (Al-Sulaiman 2014; Khatoon and Kim 2020);

$$E_j = (h_j - h_0) - T_0(h_j - s_0) \quad (10)$$

3.3. Thermal modeling of combined cycle

Heat received by the combined cycle from the SPT field is given by the heat balance equation in the heat exchanger (1);

$$\dot{Q}_r = \dot{Q}_{HX1} = \dot{m}_{ms} \cdot Cp_{ms} \cdot (h_b - h_a) = \dot{m}_{SCO2} \cdot (h_1 - h_9) \quad (11)$$

After considering the isentropic efficiency of main turbine, power output from the main turbine is determined as;

$$\dot{W}_{MT} = \dot{m}_{SCO2} \cdot (h_1 - h_{2s}) \cdot \eta_{MT} \quad (12)$$

Where, η_{MT} is the main turbine's isentropic efficiency and h_{2s} is the enthalpy at the outlet of the main turbine when expansion is done isentropically.

Heat transfer in the HTR is determined by effectiveness formula and energy balance equation;

$$\dot{Q}_{HTR} = \dot{m}_{sCO2} \cdot (h_2 - h_3) = \dot{m}_{sCO2} \cdot (h_9 - h_8) \quad (13)$$

Where, $\dot{m}_{sCO2}$ is sCO_2 mass flow rate. Effectiveness of the HTR having same mass flow rate in hot side and cold side of HTR is defined as;

$$\varepsilon_{HTR} = \frac{T_2 - T_3}{T_2 - T_8} \quad (14)$$

Power input to the pre-compressor is defined as;

$$\dot{W}_{PC} = \frac{\dot{m}_{sCO2} \cdot (h_{4s} - h_3)}{\eta_{PC}} \quad (15)$$

Where, η_{PC} is the pre-compressor isentropic efficiency. Heat transfer in LTR is also defined by applying the heat balance and effectiveness formula;

$$\dot{Q}_{LTR} = \dot{m}_{sCO2} \cdot (h_4 - h_5) = \dot{m}_{sCO2} \cdot (h_8 - h_7) \quad (16)$$

Effectiveness of the LTR is defined as;

$$\varepsilon_{LTR} = \frac{T_4 - T_3}{T_4 - T_7} \quad (17)$$

Heat absorbed by the ORC through the heat exchanger 2 (HX2) energy balance formula and the effectiveness formula can be given as;

$$\dot{Q}_{HX2} = \dot{m}_{SCO2} \cdot (h_5 - h_6) = \dot{m}_{ORC} \cdot (h_{10} - h_{13}) \quad (18)$$

Where, $\dot{m}_{ORC}$ is the ORC's working fluid mass flow rate

Effectiveness of the heat exchanger (2) is defined as;

$$\varepsilon_{LTR} = \frac{C_{ORC\cdot}(T_{10} - T_{13})}{C_{min\cdot}(T_5 - T_{13})} = \frac{C_{sCO2\cdot}(T_5 - T_6)}{C_{min\cdot}(T_5 - T_{13})} \quad (19)$$

Where, C_{ORC} and C_{sCO2} are the heat capacity of ORC working fluid and sCO_2 respectively. C_{min} is the minimum of these two values. Input power for the main compressor is defined as;

$$\dot{W}_{MC} = \frac{\dot{m}_{sCO2} \cdot (h_{7s} - h_6)}{\eta_{MC}} \quad (20)$$

Where, η_{MC} refers to isentropic efficiency of the main compressor.

Power output from the ORC turbine is determined as;

$$\dot{W}_{MT} = \dot{m}_{ORC} \cdot (h_{10} - h_{11s}) \cdot \eta_{OT} \quad (21)$$

Where, η_{OT} is the ORC turbine isentropic efficiency.

Heat rejected by the condenser can be defined as;

$$\dot{Q}_{cond} = \dot{m}_{ORC} \cdot (h_{11} - h_{12}) \quad (22)$$

Power absorbed by the pump is defined as;

$$\dot{W}_{pump} = \frac{\dot{m}_{ORC} \cdot (h_{13s} - h_{12})}{\eta_{pump}} \quad (23)$$

Where, η_{pump} is the pump isentropic efficiency.

Net power obtained from the individual bottoming ORC can be defined as;

$$\dot{W}_{net,ORC} = \dot{W}_{OT} - \dot{W}_{pump} \quad (24)$$

Furthermore, in this section exergy analysis of the combined system also to be discussed. Exergy destruction and exergy in each component are calculated by applying the exergy balance equation for each component after assuming no heat loss in the component (Cengel and Boles 2004).

Exergy equation for heat exchanger 1 is expressed as;

$$\dot{m}_{sCO2} \cdot [(h_9 - h_1) - T_0(s_9 - s_1)] + \dot{m}_{ms} \cdot [(h_a - h_b) - T_0(s_a - s_b)] - \dot{E}D_{HX1} = 0 \quad (25)$$

Exergy balance equation for the main turbine can be expressed as;

$$\dot{m}_{sCO2} \cdot [(h_1 - h_2) - T_0(s_1 - s_2)] - \dot{W}_{MT} - \dot{E}D_{MT} = 0 \quad (26)$$

For the HTR exergy balance equation can be described as;

$$\dot{m}_{sCO2} \cdot [(h_2 - h_3) - T_0(s_2 - s_3)] + \dot{m}_{sCO2} \cdot [(h_8 - h_9) - T_0(s_8 - s_9)] - \dot{E}D_{HTR} = 0 \quad (27)$$

Exergy balance equation for the pre-compressor is determined as;

$$\dot{m}_{sCO2} \cdot [(h_3 - h_4) - T_0(s_3 - s_4)] + \dot{m}_{PC} - \dot{E}D_{PC} = 0 \quad (28)$$

Exergy balance equation for LTR is given as;

$$\dot{m}_{sCO2} \cdot [(h_4 - h_5) - T_0(s_4 - s_5)] + \dot{m}_{sCO2} \cdot [(h_7 - h_8) - T_0(s_7 - s_8)] - \dot{E}D_{LTR} = 0 \quad (29)$$

Exergy balance equation for the heat exchanger (2) can be defined as;

$$\dot{m}_{sCO2} \cdot [(h_5 - h_6) - T_0(s_5 - s_6)] + \dot{m}_{ORC} \cdot [(h_{12} - h_{10}) - T_0(s_{12} - s_{10})] - \dot{E}D_{HX2} = 0 \quad (30)$$

Exergy balance equation for main compressor is expressed as;

$$\dot{m}_{sCO2} \cdot [(h_6 - h_7) - T_0(s_6 - s_7)] + \dot{m}_{MC} - \dot{E}D_{MC} = 0 \quad (31)$$

Exergy balance equation for the ORC turbine is defined as;

$$\dot{m}_{ORC} \cdot [(h_{10} - h_{11}) - T_0(s_{10} - s_{11})] - \dot{m}_{OT} - \dot{E}D_{OT} = 0 \quad (32)$$

Exergy balance equation for the condenser is defined as;

$$\dot{m}_{ORC} \cdot [(h_{12} - h_{11}) - T_0(s_{12} - s_{11})] + \dot{m}_w \cdot [(h_{15} - h_{14}) - T_0(s_{15} - s_{14})] - \dot{E}D_{cond} = 0 \quad (33)$$

Exergy balance equation for the pump is defined as;

$$\dot{m}_{sCO2} \cdot [(h_{13} - h_{12}) - T_0(s_{13} - s_{12})] + \dot{W}_{pump} - \dot{E}D_{pump} = 0 \quad (34)$$

After calculating the exergy destruction rate for each component, total exergy destruction rate for the combined cycle is calculated as;

$$\dot{E}D_{total} = \dot{E}D_{HX1} + \dot{E}D_{MT} + \dot{E}D_{HTR} + \dot{E}D_{PC} + \dot{E}D_{LTR} + \dot{E}D_{HX2} + \dot{E}D_{MC} + \dot{E}D_{OT} + \dot{E}D_{cond} + \dot{E}D_{pump} \quad (35)$$

3.4. Performance parameters for the proposed model

On the basis of the thermal modeling, numerous mathematical relations are used in the performance analysis of the solar power tower powered combined cycle have been discussed below;

Net power output obtained from the combined cycle is defined as;

$$\dot{W}_{net} = \dot{W}_{MT} + \dot{W}_{OT} - \dot{W}_{PC} - \dot{W}_{MC} - \dot{W}_{pump} \quad (36)$$

Solar-powered combined cycle's thermal efficiency is determined as;

$$\eta_{th} = \frac{\dot{W}_{net}}{\dot{Q}_{solar}} \quad (37)$$

Combined cycle exergy efficiency is determined as (Al-Sulaiman 2014; Cengel and Boles 2004; Khan and Mishra 2020)

$$\eta_{ex} = 1 - \frac{\dot{E}D_{total}}{\dot{E}_{solar}} \quad (38)$$

The thermal efficiency of the combined cycle can also be defined by the relation between thermal and exergy efficiency of the combined cycle (Cengel and Boles 2004; Khan and Mishra 2020);

$$\eta_{th} = \eta_{ex} \cdot \eta_{Carnot} \quad (39)$$

At last a performance parameter is to be defined which represents the capacity to recover waste heat from the topping cycle. That parameter is waste heat recovery ratio (WHRR). WHRR is defined for the bottoming cycle. WHRR is defined as the ratio of the net power output (net power output of ORC) to the maximum available waste heat to be recovered from waste heat source (Kim et al. 2016b). WHRR for the bottoming cycle is defined as;

$$WHRR = \frac{\dot{W}_{net,ORC}}{\dot{m}_{sCO2} \cdot (h_5 - h_0)} \quad (40)$$

Where, h_0 is the enthalpy of waste heat of the topping cycle at environmental temperature.

Modeling equations of the SPT powered combined cycle were solved in engineering equation solver (EES) (Klein 2019).

3.5. Selection for working fluids

Care must be taken when selecting the working fluid for any thermodynamic cycle because it affects the cycle performance, economic feasibility, and environmental aspects (Wang et al. 2019). A mixture of magnesium dichloride (MgC_{l2}) and potassium chloride has been used as molten salt HTF in the receiver with mass fraction of 32 and 68%, respectively (Khatoon and Kim 2020). Reason behind choosing this HTF is that this is the cheapest option for the heliostat-driven sCO_2 cycle as compared to the solar salt and liquid sodium (Na) (Polimeni et al. 2018). Thermo-physical properties of this molten salt are listed in Table 2. It is challenging to select working fluid for the ORC because it loses its chemical stability above its maximum temperature therefore it obtains optimum thermo-physical properties at optimum pressure and temperature (Koc, Yagli, and Koc 2019). Working fluids for the ORC system are categorized as dry, isentropic, and wet fluid. Dry and isentropic working is more suitable than the other type of fluids due to high-quality vapor at the expander outlet (Saleh et al. 2019). Also in current study waste heat source has low temperature. Therefore, due to these reasons and low-temperature applications, five working fluids such as isopentane, R245fa, R236fa, isobutene, and R227ea were considered for the analysis of ORC in this study. Thermal properties, safety, and environmental data of these working fluids are given in Table 3. The safety group classification for each refrigerant has two or three alphanumeric characters (e.g., B1 or A2L). The first character shows the toxicity and the numeral, with or without a suffix letter, indicates the flammability. There are two classes for toxicity: lower toxicity (Class A) and higher toxicity (Class B). There are four classes of flammability: 1, 2 L, 2 or 3(Calm 1994).

3.6. Validation of the proposed model

In order to ensure the correct use of the modeling equation, previous studies were used to validate the current model. There is no availability of the literature on combined pre-compression sCO_2 cycle and ORC. Therefore, both toping and bottoming cycles were validated separately with existing literature. Thermal efficiency is considered as the validation parameter for both the cycles. Pre-compression sCO_2 cycle was validated with the previous studies Kim et al. (2016a); Yu et al. (2020) at same input conditions. Apart from this, bottoming ORC was also validated with previous study Clemente et al. (2013) at same input conditions. Thermal efficiency of the both topping and bottoming cycle was found very close to the respective previous studies as shown in Tables 4 and 5 respectively. However, present study was performed with input parameters that are different from the previous studies those were used for the validation purpose.

Table 2. Thermo-physical properties of molten salt (magnesium dichloride + potassium chloride) (Xu et al. 2018).

Parameters	Values
Density	1593 (kg/m^3)
Specific heat capacity	1.028 (kJ/kg-K)
Thermal conductivity	0.39 (W/m-K)
Solidification temperature	699 K
Stability limit	1691 K

4. Results and discussion

Parametric analysis of the solar power tower integrated combined pre-compression sCO_2 cycle and ORC is conducted in current study. A computer program was made in the engineering equation solver (EES) to simulate the combined cycle's performance based on the assumptions in this study. Further, calculated thermodynamic properties at main state points are shown in Table 6.

4.1. Comparison with standalone cycle

Firstly, combined cycle performance was compared to the standalone pre-compression cycle under standard conditions. Based on assumed input data, such as solar irradiation at 970 W/m^2, maximum cycle temperature, and pressure at 650°C and 25 MPa, main inlet compressor pressure and temperature at 7.5 MPa (Yu et al. 2020) and 32°C, combined cycle thermal efficiency and net power output were achieved at 46.67% and 254.4 kW, respectively, based on R245fa working fluid. While at similar input conditions, the standalone cycle yielded 44.65% and 243.4 kW of thermal efficiency and output power respectively. This means, application of organic Rankine cycle to standalone pre-compression cycle as bottoming cycle improved thermal efficiency and the power output, respectively, by 4.52 and 4.51%.

Furthermore, comparison of the performance was examined with the variation of solar irradiation. Thermal efficiency and power output of the standalone cycle and the combined cycle have the same trend as can be shown in Figures 3 and 4 respectively. As solar irradiation increased from 500 to 900 W/m^2, thermal efficiency of the standalone and combined cycle improved was by 97.3 and 75.9% respectively.

Table 3. Properties of organic working fluids (Saleh et al. 2019; Saleh and Aly 2017).

Working substance	P_c (MPa)	T_c (°C)	T_b* (°C)	Weight (Kg/Kmole)	Type	ODP	GWP	Lifetime (years)	Security group
Isopentane	3.38	187.2	27.8	72.1	Dry	0	20	0.009	A3
R245fa	3.65	154,1	15.1	134.05	Dry	0	1050	7.7	B1
R236fa	3.2	124.9	−1.5	152	Dry	0	9890	242	A1
Isobutene	3.63	134.7	−11.7	58.1	Dry	0	20	0.016	A3
R227ea	2.93	101.8	−16.3	170.03	Dry	0	3580	38.9	A1

* T_b is corresponding to the atmospheric pressure

Table 4. Validation of toping pre-compression sCO_2 cycle.

Input parameters	Kim et al. 2016a	Yu et al.2020
P_1 (MPa)	27.18	25
T_1 (°C)	384.85	500
P_6 (MPa)	8.82	7.5
T_6 (°C)	36.85	32
η_{MC}	0.85	0.85
η_{MT}	0.9	0.88
Reference thermal efficiency (%)	31.42	37.88
Current model thermal efficiency (%)	31.6	39.85
Estimated errors in thermal efficiency (%)	0.05	4.48

Table 5. Validation of bottoming ORC.

Input parameters	Clemente et al. 2013 with Isopentane	Clemente et al. 2013 with R245fa
T_{10} (°C)	184.1	154.2
P_{10} (MPa)	3.023	3.395
P_{11} (MPa)	0.1515	0.2504
η_{OT}	0.6	0.6
η_{pump}	0.5	0.5
Reference thermal efficiency	0.12	0.11
Current model thermal efficiency	0.129	0.116
Estimated errors in thermal efficiency (%)	7.5	5.45

Table 6. Thermodynamic properties at main state points.

State points	Working fluids	Temperature (°C)	Pressure (MPa)	Enthalpy (kJ/kg)	Entropy (kJ/kg-K)
1	sCO_2	650	25	651.1	0.09878
2	sCO_2	465.3	5.435	440.1	0.3691
3	sCO_2	119.6	5.6	53.07	−0.5719
4	sCO_2	129.9	6.216	61.27	−0.5688
5	sCO_2	97.63	6.216	23.42	−0.6667
6	sCO_2	32	7.5	−144.1	−1.206
7	sCO_2	95.73	23.48	−106.2	−1.1916
8	sCO_2	112.5	25	−75.42	−1.117
9	sCO_2	360.1	25	287.7	−0.3755
10	R227ea	76.5	3	113.4	0.3691
11	R227ea	30.69	0.1504	101.2	0.3805
12	R227ea	25	0.1504	46.6	0.1751
13	R227ea	27.07	3	51.24	0.1777

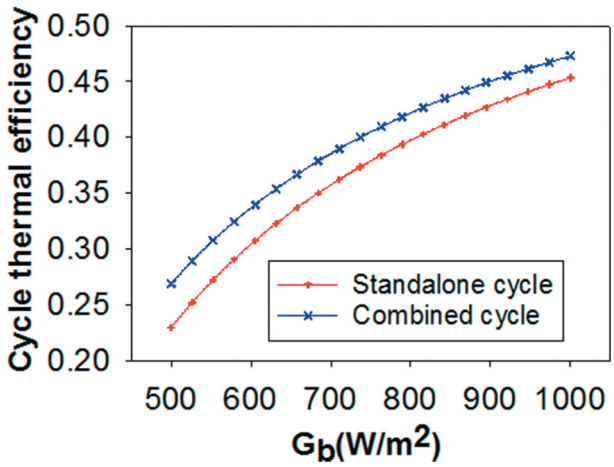


Figure 3. Thermal efficiency comparison of cycles at varying solar irradiation.

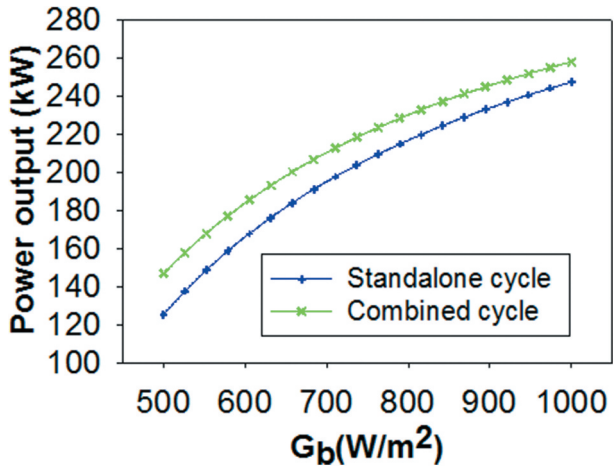


Figure 4. Power output comparison of both cycles with varying solar irradiation.

4.2. Effect of solar irradiation on system performance

As present combined model is powered by solar power tower therefore it is necessary to examine the effect of solar irradiation on the system performance. Exergy efficiency of the combined cycle was increased with the solar irradiation continuously. Reason behind this is that increased solar irradiation is effectively utilized by the solar concentrator field. This leads to increase inlet exergy to the combined cycle (Khan and Mishra 2020). It can also be explained as solar irradiation increases, solar exergy inlet to the combined cycle also increases according to the Eq. (7). However, the rate of total exergy destruction with the solar irradiation was not affected. Therefore, the exergy efficiency was improved according to the Eq. (38), as can be seen in Figure 5. As solar irradiation changed from 500 to 1000 W/m^2, exergy efficiency changed from 44.11 to 74.06% on the basis of R227ea working fluid. Isopentane provided the lowest exergy efficiency, among other considered working fluids.

As can be seen from Figure 6, the thermal efficiency of the combined cycle increased with the solar irradiation. R227ea gave the highest thermal efficiency among the other selected working fluids. Highest thermal efficiency was improved by 65.05% as solar irradiation increased from 500 to 1000 W/m^2.

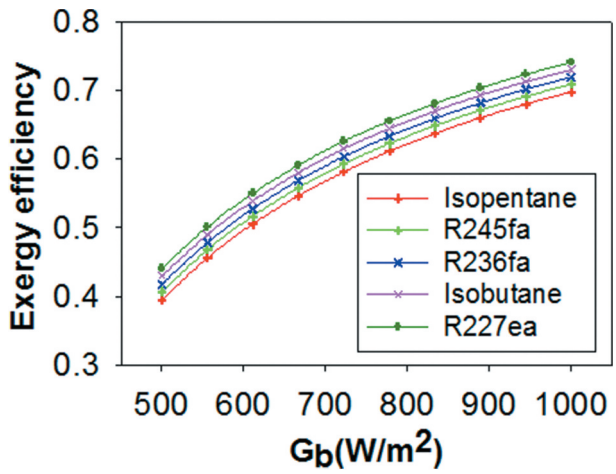


Figure 5. Variation of the combined-cycle exergy efficiency with solar irradiation.

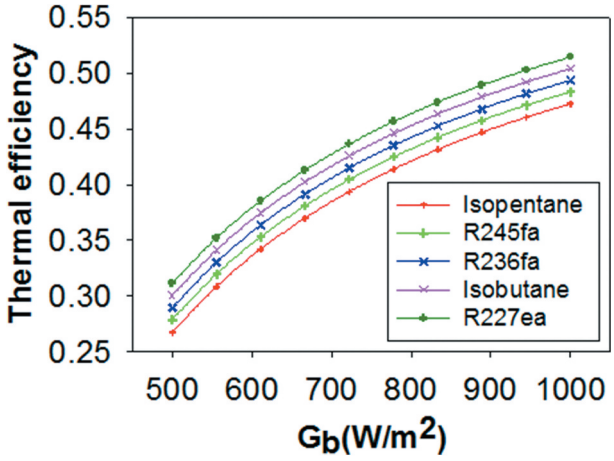


Figure 6. Variation of combined cycle-thermal efficiency with solar irradiation.

Thermal efficiency has the direct relation with the exergy efficiency (Cengel and Boles 2004). Therefore, as exergy efficiency increased thermal efficiency also increased with solar irradiation according to Eq. (39).

Apart from this, power output of the combine cycle also increased with solar irradiation as shown in Figure 7. Increased solar irradiation increased the enthalpy at the turbine inlet. This result, increase in turbine output work. This leads to increase in total output power. Highest output power was measured 278.5 kW at 1000 W/m^2 solar irradiation using R227ea.

4.3. Effect of main turbine inlet temperature on system performance

The temperature of the main inlet turbine (MTIT) is directly affected by the temperature of the molten salt. MTIT increases with the temperature of the molten salt. However, the increased temperature of the molten salt will increase the heat loss in the solar receiver. This results in lower receiver efficiency (Khatoon and Kim 2020). However, in the present study, the parameters of the solar field are kept constant and shown in Table 1. The main objective is therefore to investigate the combined cycle performance. The effects of MTIT on system performance are shown in Figures 8–10. While investigating the effects of MTIT, other input parameters have been kept constant as shown in Table 1. Figures 8–10 show that exergy, thermal efficiency, and power output have increased with MTIT. The reason behind increased efficiency and power output is that as MTIT increases the enthalpy difference across the MT, resulting in increased output power. However, efficiency and output power have a different trend as MTIT increases, as shown in Figures 8–10. Highest exergy and thermal efficiency and power output were achieved 72.9, 52.7% and 322.7 kW, respectively, at 800 °C using R227ea. Efficiency and output power are therefore highest with higher input thermodynamic variables. However, higher thermal performance cannot be achieved due to safety and material limitations. Although at same solar irradiation supercritical carbon dioxide cycle achieved higher efficiency than the superheated steam cycle at higher input turbine conditions (Chacartegui et al. 2011)

4.4. Effects of main compressor inlet pressure on system performance

During the investigation of the effect of MCIP, MC inlet temperature, MT inlet pressure and temperature, PC inlet pressure was fixed at 32 °C, 25 MPa, 650°C and 5.6 MPa, respectively. It is found that in the subcritical region where the pressure is below the critical value of 7.38 MPa the thermal and exergy efficiency increases with MCIP. Beyond the critical pressure of CO_2, thermal and exergy efficiency increase and thereafter decrease. This means that there is an optimum pressure where net power output and efficiency have reached their highest values. Exergy and thermal efficiency and net output power of the system have a bell-shaped curve as can be seen in Figures 11–13 respectively. The reason behind this pattern is that the highest density of CO_2 at critical pressure leads to lower compression power (Blanco and Santigosa 2017). This results in the highest net power output and consequently the highest thermal and exergy efficiency. Using R227ea, the highest thermal and exergy efficiency and output power were achieved by 50.33, 74.34% and 274.3 kW at an optimum pressure of 7.737 MPa. Figures 11–13, also depicts that the efficiencies and net output power have same trend due

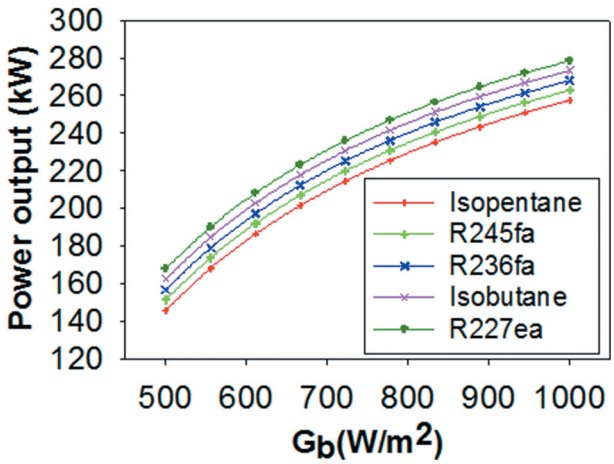


Figure 7. Variation of the combined cycle power output with solar irradiation.

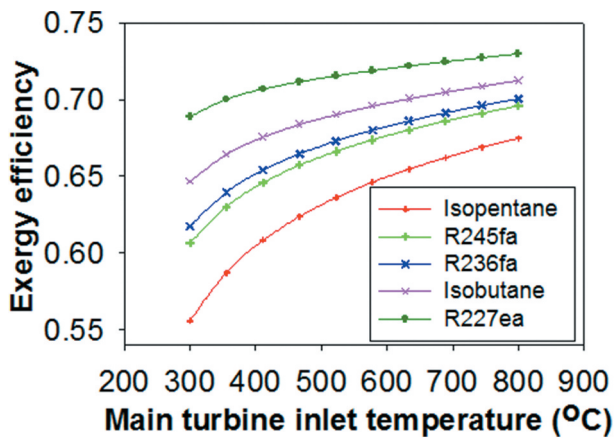


Figure 8. Variation of exergy efficiency with the main turbine inlet temperature.

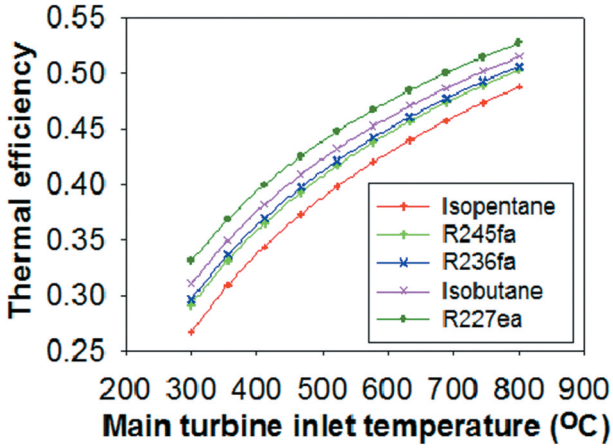


Figure 9. Variation of thermal efficiency with main turbine inlet temperature.

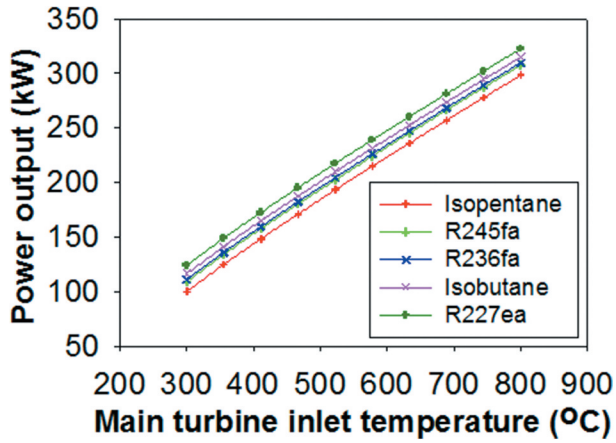


Figure 10. Variation in power output with main turbine inlet temperature.

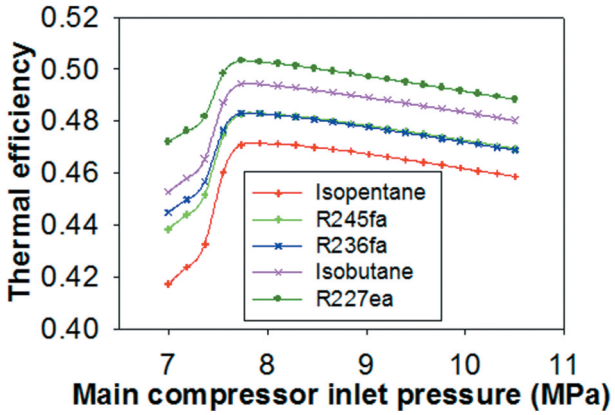


Figure 11. Variation of the combined cycle- thermal efficiency with main compressor inlet pressure.

to heat available at the inlet of the combined cycle is fixed and only varies with solar field outlet temperature (Yu et al. 2020). Additionally, performance of the system almost overlaps with working fluids R245fa and R236fa. This is because of the both fluids almost having similarity in thermophysical properties as listed in Table 3.

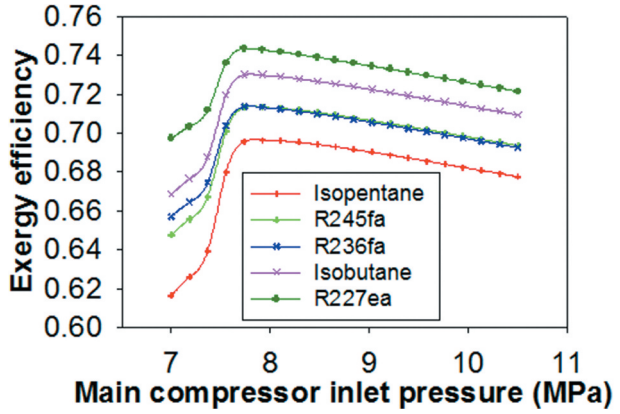


Figure 12. Variation of the combined cycle exergy efficiency with the main compressor inlet pressure.

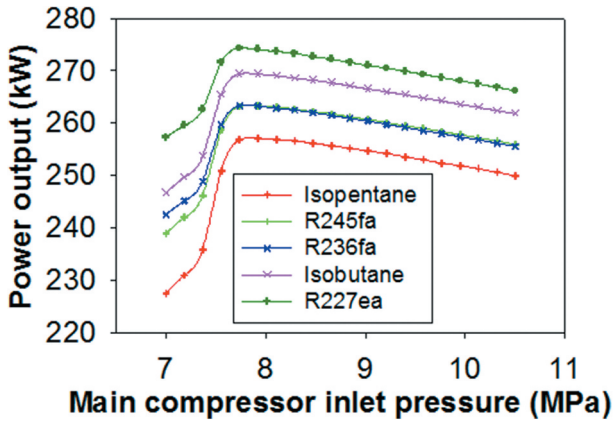


Figure 13. Variation of power output of the combined cycle with the main compressor inlet pressure.

4.5. Effects of main compressor inlet temperature on system performance

System performance is also influenced by the main compressor inlet temperature (MCIT). The influence of the MCIT on system performance was investigated by keeping the other input parameters constant as shown in Table 1. The MCIP has been fixed at 7.5 MPa. The MCIT ranges from 32 to 38 °C Exergy and thermal efficiency and net output of the combined system were decreased with the MCIT as shown in Figures 14–16. The reason behind the decrease in thermal performance of the system is that, as the temperature rises above the critical state of CO_2, the specific heat of the sCO_2 decreases, which leads to a lower enthalpy at the MC inlet (Chen and Lundqvist 2011). This means more enthalpy difference across the compressor, which leads to more input power. This results in a lower net output power. The heat available at the inlet to the combined cycle is known to be constant, depending on the SPT parameters. As a result, the efficiency and net output of the combined cycle were reduced with the MCIT. It was also shown that the total decrease in thermal performance of the combined system was due to the standalone pre-compression cycle. Because the performance of the bottoming ORC was not affected by the variation of the MCIT. Highest exergy and thermal efficiency and power output were obtained using

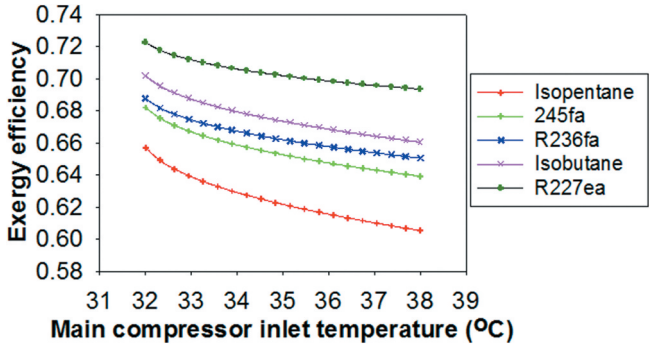


Figure 14. Variation of the exergy efficiency with main compressor inlet temperature.

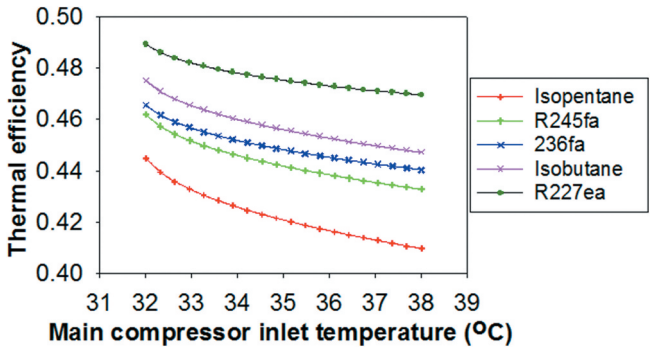


Figure 15. Variation of the thermal efficiency with main compressor inlet temperature.

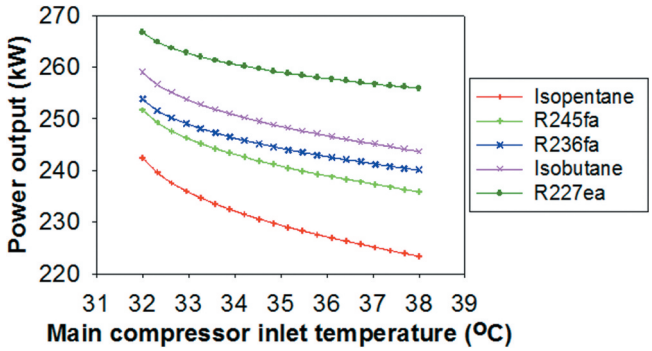


Figure 16. Variation of the power output with main compressor inlet temperature.

R227ea having values 72.26%, 48.92% and 266.7 kW, respectively, at 32 °C MCIT. On other hand isopentane gave lowest values.

4.6. Effect of inlet pressure of pre-compressor on system performance

Since the present study is all about the thermal performance investigation of the combined pre-compression cycle of sCO_2 and ORC. It is therefore necessary to investigate the effects of the pre-compressor inlet pressure (PCIP) on the combined system performance. The effect of PCIP was therefore examined keeping other parameters constant as shown in Table 1. The performance of the system was investigated between 5.6 and 6.8 MPa PCIP. Figures 17–19 show that the exergy and thermal efficiency and net output power of the combined system increased with the PCIP. Because the at low temperature, density of carbon dioxide increases with an increase in the pressure. This leads to decrease in PC input power. Therefore, net power output of the combined cycle increases. This leads to increases in thermal and exergy efficiency of the system. From Figure 17, it can be seen that using the R227ea, exergy efficiency was improved from 72.26 to 74.45%, when PCIP increased from 5.6 to 6.8 MPa. While thermal efficiency and the maximum net power output were found to be increased from 48.92 to 50.4% and 266.7 to 269.3 kW, respectively, as PCIP increased from 5.6 to 6.8 MPa with same working fluid R227ea as shown in Figures 18 and 19 respectively.

It can be seen that R227ea has the highest combined cycle performance among the other selected working fluids due to its thermo-physical properties. Its critical temperature and pressure are below the maximum ORC temperature as shown in Table 3. This means that R227ea is working under a super critical state. As is known, the system delivers the highest performance when operating fluid is near or above critical condition (Kalra et al. 2012).

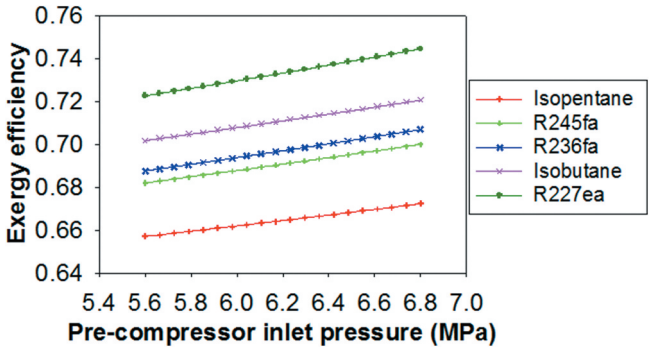


Figure 17. Variation of exergy efficiency of the combined system with pre-compressor inlet pressure.

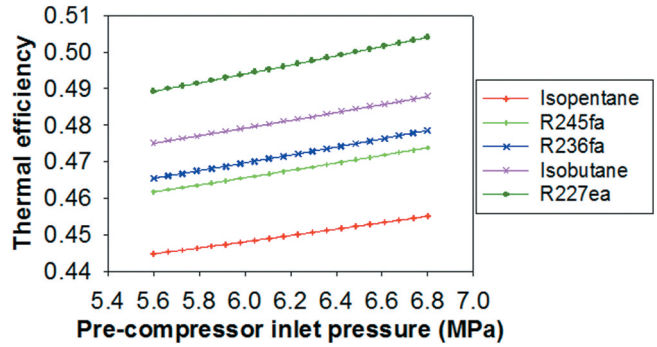


Figure 18. Variation of the thermal of the combined system with the pre-compressor inlet pressure.

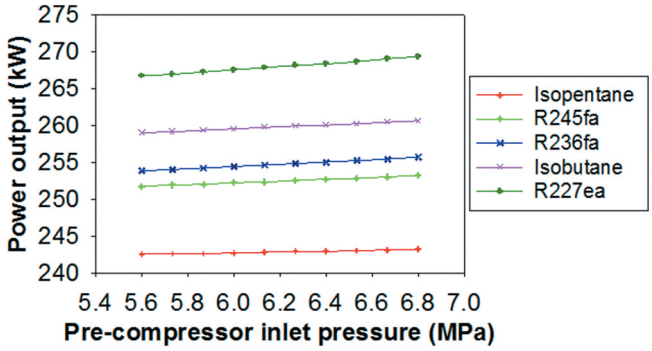


Figure 19. Variation of the combined cycle- power output with pre-compressor inlet pressure.

4.7. Effect of the HX2 effectiveness on system performance

It is also necessary to examine the effect of the HX2 effectiveness on the performance of the system, since HX2 is the main component for the transfer of heat from the topping cycle to the bottoming cycle. As can be seen in Figure 20, the thermal efficiency and net power output of the combined system have increased continuously with HX2 effectiveness. The reason for this increased efficiency and power output is that more heat is recovered in the HX2 with effectiveness. As can be seen from Figure 20, although the performance of the system was continuously improved with effectiveness, the maximum effectiveness was taken as 0.95 for the design of the heat exchanger because, above that value, the material cost of the heat exchanger increased (Abdullah, Al-Zahrani, and Dincer 2018).

In addition, Figure 20 is based on the working fluid of R227ea. The system's highest thermal efficiency and power output increased from 48.66 to 49.2% and 265.3 to 268.2 kW, respectively, when efficiency increased from 0.85 to 0.95. As noted, there has been a slight change in the combined system performance. Because the bottoming cycle performance did not contribute much to the combined cycle performance as explained in the previous section of this study.

4.8. Effects of system variable on waste heat recovery ratio

The key parameter to define the performance of the bottoming cycle is the waste heat recovery ratio as defined above. It has already been established in this study that the addition of the bottoming cycle to the basic pre-compression cycle is the objective of the present study. It is therefore necessary to examine a parameter that can be used to define the utility of the ORC application in the topping cycle. That's why the ratio of waste heat recovery has been defined. HX2 is responsible for the thermal performance of the bottoming cycle. Therefore, variation of WHRR with HX2 effectiveness with different working fluids is shown in Figure 21. WHRR increased continuously with the effectiveness of HX2. The increase in WHRR with HX2 effectiveness can be explained as the recovery of waste heat in HX2 is increased with effectiveness as explained in the above section. As a result, more enthalpy at the inlet of the ORC turbine leads to increased work output across the turbine. As a result, WHRR increased according to the Eq. (40).

Maximum WHRR at effectiveness 0.95 was found to be 0.5673, 0.3394, 0.2016 and 0.1172 using R227ea, isobutene, R236fa, R245fa and Isopentane, respectively. It has also been observed that all working fluids have different WHRR improvement rates. As effectiveness increased from 0.85 to 0.95, WHHR improved by 16.41, 16.39, 39.13, 40.73, and 52.60% for R227ea, isobutene, R236fa, R245fa and Isopentane, respectively. This means that it is of major concern to select the proper working fluid for the bottoming ORC for the efficient recovery of waste heat from the topping cycle.

Further in this segment, the effectiveness of LTR affects the WHRR. Effect of the effectiveness of the LTR on the performance of the ORC was also examined. WHRR was slightly increased with the effectiveness of the LTR as shown in Figure 22. The reason behind this increased WHRR is that when the effectiveness of the LTR is increased, more heat is recovered by the cold stream of the sCO_2. This leads to a lower sCO_2 temperature at inlet to HX2. In other words, the low heat at the inlet to the HX2 can be said. It reduces the inlet temperature of the ORC turbine. As a result, the lower inlet temperature of the ORC turbine increases the output power of the ORC turbine in case of organic working fluids (Dai, Wang, and Gao 2009). On the other hand at state 5 enthalpy also decreases because of lower heat at the HX2 inlet. As a result, WHRR increased with the LTR effectiveness according to the Eq. (40). As a consequence of the study, the highest WHRR value was found to be 0.56, 0.3352, 0.1775, 0.1197 and 0.0998 using R227ea, isobutene, R236fa, R245fa and isopentane, respectively, as shown in Figure 22.

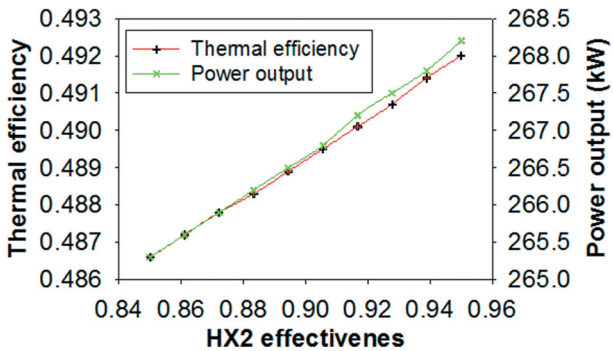


Figure 20. Variation of the combined-cycle thermal efficiency and power output with heat exchanger 2 effectiveness.

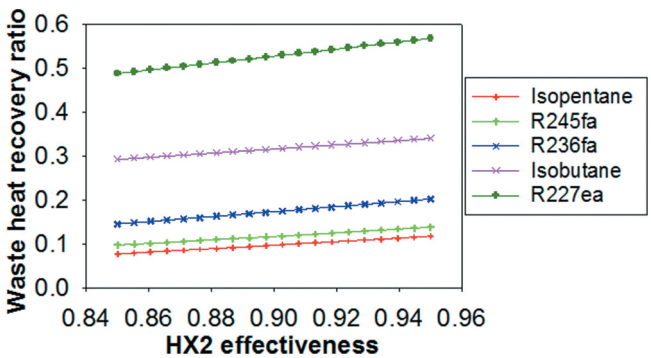


Figure 21. Variation of the waste heat recovery ratio with the heat exchanger 2 effectiveness.

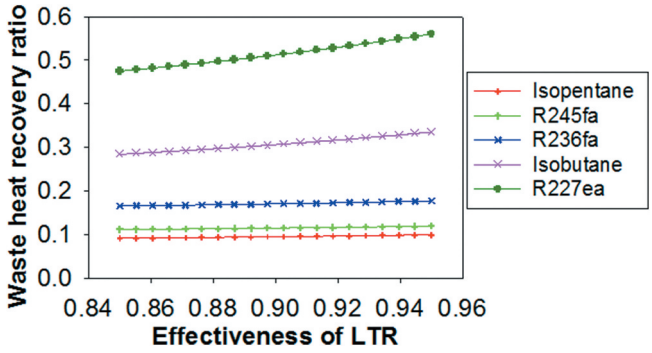


Figure 22. Variation of the waste heat recovery ratio with the LTR effectiveness.

5. Conclusions

Thermodynamic analysis on the SPT powered combined pre-compression sCO_2 cycle and ORC was conducted in this study. Based on the results following conclusions were made;

- Thermal efficiency and output power of the standalone pre-compression cycle improved by 4.52% and 4.51%, respectively, by recovering waste heat using bottoming ORC.
- The net power output and efficiency of the combined cycle increased continuously with solar irradiation, main turbine inlet temperature, inlet pressure of pre-compressor and heat exchanger 2 effectiveness.
- Highest exergy, thermal efficiency, and net power output were found to be 74.06, 51.83%, and 278.5 kW, respectively, at 1000 W/m^2 of solar irradiation using R227ea.
- Exergy, thermal efficiency, and net power output increased first, and then decreased with inlet pressure of main compressor. Using R227ea, best performance with optimum pressure of 7.737MPa was found.
- Efficiency and the power output of combined cycle was decreased with main compressor inlet temperature and was highest at 32 °C.
- The waste heat recovery ratio was continuously increased with the efficiency of the heat exchanger 2 and LTR. Its rate of improvement was also different for different working fluids. The highest values were 0.5673 at 0.95 heat exchanger 2 effectiveness and 0.56 at 0.95 LTR effectiveness.
- In order to achieve better performance, proper working fluid selection is needed. R227ea and Isobutane are recommended as the best and worst working fluids among the other fluids considered in this study, respectively.
- In the future study, a detailed study on the selection of fluids to be carried out along with an economic analysis of the current model.

Nomenclature

A_h	Area of single heliostat (m^2)
Cp	constant pressure specific heat (kJ/kg-K)
$\dot{E}D$	Exergy destruction rate (kW)
$\dot{E}_{solar}$	Solar exergy inlet to combined cycle (kW)
$\dot{E}$	Exergy rate (kW)
f_{view}	View factor of receiver
G_b	Solar irradiation (W/m^2)
h_{conv}	Convective heat loss coefficient (W/m^2-K)
h	specific enthalpy (kJ/kg)
N_h	Number of heliostat
$\dot{m}$	Mass flow rate (kg/s)
$\dot{Q}$	Heat rate in (kW)
$\dot{Q}_h$	Actual solar heat received by heliostat field (kW)
$\dot{Q}_{loss,r}$	Heat loss from the receiver (kW)
$\dot{Q}_r$	Heat received by central receiver (kW)
$\dot{Q}_{solar}$	Solar heat received by heliostat field (kW)
sCO_2	Supercritical carbon dioxide
s	specific entropy (kJ/kg-K)
T_R	Receiver surface temperature (K)
T	Temperature (K)
$\dot{W}$	Power (kW)
η_h	Heliostat efficiency
η_r	Receiver thermal efficiency
η_{ex}	Exergy efficiency
η_{th}	Thermal efficiency

Abbreviations

CR	concentration ratio
Cond	condenser
HX1	heat exchanger 1
HTR	High temperature recuperator
HX2	heat exchanger 2
LTR	Low temperature recuperator
MC	Main compressor
MCIP	Main compressor inlet pressure (MPa)
MCIT	Main turbine inlet temperature (°C)
MT	Main turbine
MTIT	Main turbine inlet temperature (°C)
OT	ORC turbine\
ORC	Organic Rankine cycle
PC	Pre-compressor
PCIP	Pre-compressor inlet pressure (MPa)
SPT	solar power tower
WHRR	Waste heat recovery ratio

Greek letters

α	solar absorbance
η	efficiency
σ	Stephen Boltzmann constant (W/m)
ε	effectiveness
β	Sun's subtended cone half angle(rad)
δ	Change in property
ζ	thermal emittance

Subscripts

ms	molten salt
i	inlet
e	exit
0	environmental conditions
h	heliostat
r	receiver
b	boiling
c	critical

References

Abdullah, A., Al-Zahrani, and I. Dincer. 2018. Energy and exergy analyses of a parabolic trough solar power plant using carbon dioxide power cycle. *Energy Conversion and Management* 158:476–88. doi:10.1016/j.enconman.2017.12.071.

Abid, M., V. O. Adebayo, and U. Atikol. 2019. Energetic and exegetic analysis of a novel multi-generation system using solar power tower. *International Journal of Exergy* 29:211. doi:10.1504/IJEX.2019.100364.

Ahn, Y., S. J. Bae, M. Kim, S. K. Cho, S. Baik, J. I. Lee, and J. E. Cha. 2015. Review of supercritical CO2 power cycle technology and current status of research and development. *Nuclear Engineering Technology* 47:647–61. doi:10.1016/j.net.2015.06.009.

Alsagri, A. S., A. Chiasson, and M. Gadalla. 2018. Viability assessment of a concentrated solar power tower with a supercritical CO2 Brayton cycle power plant. *Journal of Solar Energy Engineering*. doi:10.1115/1.4043515.

Al-Sulaiman, F. A. 2014. Exergy analysis of parabolic trough solar collectors integrated with combined steam and organic Rankine cycles. *Energy Conversion and Management* 77:441–49. doi:10.1016/j.enconman.2013.10.013.

Al-Sulaiman, F. A., and M. Atif. 2015. Performance comparison of different supercritical carbon dioxide Brayton cycles integrated with a solar power tower. *Energy* 82:61–71. doi:10.1016/j.energy.2014.12.070.

Barlev, D., R. Vidu, and P. Stroeve. 2011. Innovation in concentrated solar power. *Solar Energy Materials and Solar Cells* 95 (10):2703–25. doi:10.1016/j.solmat.2011.05.020.

Bejan, A., D. W. Kearney, and F. Kreith. 1981. Second law analysis and synthesis of solar collector systems. *Journal of Solar Energy Engineering, Transactions of the ASME* 103:23–28. doi:10.1115/1.3266200.

Besarati, S. M., and D. Y. Goswami. 2014. Analysis of advanced supercritical carbon dioxide power cycles with a Bottoming cycle for concentrating solar power applications. *Journal of Solar Energy Engineering* 136:010904-1-7. doi:10.1115/1.4025700.

Bishoyi, D., and K. Sudhakar. 2017. Modeling and performance simulation of 100 MW PTC based solar thermal power plant in Udaipur India. *Case Studies in Thermal Engineering* 10:216–26. doi:10.1016/j.csite.2017.05.005.

Blanco, M., and L. R. Santigosa. 2017. *Advances in concentrating solar thermal research and technology*. Woodhead publishing series in energy (Elsevier) Kidligton, United Kingdom. ISBN: 978-0-08-100516-3.

Calm, J. M. 1994. Referegerant safety. *ASHRAE Journal* 36 (7):17–26.

Cengel, Y. A., and M. A. Boles. 2004. *Thermodynamics an engineering approach*. 5th ed. New York, USA: McGraw-Hill publication.

Chacartegui, R., J. M. Muñoz, D. Escalona, D. Sánchez, B. Monje, and T. Sánchez. 2011. Alternative cycles based on carbon dioxide for central receiver solar power plants. *Applied Thermal Engineering* 31:872–79. doi:10.1016/j.applthermaleng.2010.11.008.

Chen, Y., and P. Lundqvist. 2011. The CO2 transcritical power cycle for low grade heat reconery discussion on temperature profiles in system heat exchangers. *Proceedings of the ASME Power Conference POWER* July 12–14, Denver, Colorado, USA.

Clemente, S., D. Micheli, M. Reini, and R. Taccani. 2013. Bottoming organic Rankine cycle for a small scale gas turbine: A comparison of different solutions. *Applied Energy* 106:355–64. doi:10.1016/j.apenergy.2013.02.004.

Dai, Y., J. Wang, and L. Gao. 2009. Parametric optimization and comparative study of organic Rankine cycle (ORC) for low grade waste heat recovery. *Energy Conversion and Management* 50:576–82. doi:10.1016/j.enconman.2008.10.018.

Dostal, V. 2004.A supercritical carbon dioxide cycle for next generation nuclear reactors. *Thesis submitted in Massachusetts institute of technology.USA*

Farges, O., J. Bézian, and M. El-Hafi. 2018. Global optimization of solar power tower systems using a Monte Carlo algorithm: Application to a redesign of the PS10 solar thermal power plant. *Renewable Energy* 119:345–53. doi:10.1016/j.renene.2017.12.028.hal-01660563.

Fuqianga, W., C. Ziminga, T. Jianyua, Y. Yuanc, S. Yongc, and L. Linhua. 2017. Progress in concentrated solar power technology with parabolic trough collector system: A comprehensive review. *Renewable and Sustainable Energy Reviews* 79:1314–28. doi:10.1016/j.rser.2017.05.174.

Guo, C., D. D. Xiaoze, Y. Goswami, and L. Yang. 2016. Investigation on working fluids selection for organic rankine cycles with low-temperature heat sources. *International Journal of Green Energy* 13 (6):556–65. doi:10.1080/15435075.2014.979491.

Ho, C. K., and B. D. Iverson. 2014. Review of high-temperature central receiver designs for concentrating solar power. *Renewable Sustainable Energy Reviews* 29:835–46. doi:10.1016/j.rser.2013.08.099.

Hoang, A. T. 2018. Waste heat recovery from diesel engines based on Organic Rankine Cycle. *Applied Energy* 231:138–66. doi:10.1016/j.apenergy.2018.09.022.

Kabira, E., P. Kumar, S. Kumar, A. A. Adelodund, and K. Kime. 2018. Solar energy: Potential and future prospects. *Renewable and Sustainable Energy Reviews* 82:894–900. doi:10.1016/j.rser.2017.09.094.

Kalogirou, S. A. 2004. Solar thermal collectors and applications. *Progress in Energy and Combustion Science* 30 (3):231–95.

Kalra, C., G. Becquin, J. Jackson, A. L. Laursen, H. Chen, K. Myers, A. Hardy, H. Klockow, and J. Zia.2012.High-potential power cycles & working fluids for next generation binary supercritical organic Rankine cycle for enhanced geothermal systems. *Proceedings, Thirty-Seventh Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, California.

Khan, Y., and R. S. Mishra. 2020. Parametric (exergy-energy) analysis of parabolic trough solar collector-driven combined partial heating supercritical CO2 cycle and organic Rankine cycle. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. doi:10.1080/15567036.2020.1788676.

Khatoon, S., and M. Kim. 2020. Performance analysis of carbon dioxide based combined power cycle for concentrating solar power. *Energy Conversion and Management* 20. doi:10.1016/j.enconman.2019.112416.

Kim, M. S., Y. Ahn, B. Kim, and J. I. Lee. 2016a. Study on the supercritical CO2 power cycles for landfill gas firing gas turbine bottoming cycle. *Energy* 111:893–99. doi:10.1016/j.energy.2016.06.014.

Kim, Y. M., D. G. Shin, C. G. Kim, and G. B. Cho. 2016b. Single-loop organic Rankine cycles for engine waste heat recovery using both low- and high-temperature heat sources. *Energy* 96:482–94. doi:10.1016/j.energy.2015.12.092.

Klein, S. A.,2019. Engineering equation solver (EES), *Academic Commercial V7.714. F-Chart Software*. www.fChart.com.

Koc, Y., H. Yagli, and A. Koc. 2019. Exergy analysis and performance improvement of a subcritical/supercritical Organic Rankine Cycle (ORC) for exhaust gas waste heat recovery in a biogas fuelled Combined Heat and Power (CHP) 520 engine through the use of regeneration. *Energies* 12 (4):575. doi:10.3390/en12040575.

Kulhanek, M., and V. Dostal. 2011 Supercritical carbon dioxide cycles thermodynamic analysis and comparison I. Proc. Scco2 Power Cycle Symp, Prague. 2011.

Li, J., Z. Ge, Y. Duan, Z. Yang, and Q. Liu. 2018. Parametric optimization and thermodynamic performance comparison of single-pressure and dual-pressure evaporation organic Rankine cycles. *Applied Energy* 217:409–21. doi:10.1016/j.apenergy.2018.02.096.

Maa, Y., T. Morozyukb, M. Liuc, J. Yanc, and J. Liua. 2019. Optimal integration of recompression supercritical CO2 Brayton cycle with main compression intercooling in solar power tower system based on exergoeconomic approach. *Applied Energy* 242:1134–54. doi:10.1016/j.apenergy.2019.03.155.

Maali, R., and T. Khir. 2020. Performance analysis of different orc power plant configurations using solar and geothermal heat sources. *International Journal of Green Energy* 17:349–62. doi:10.1080/15435075.2020.1731517.

Neises, T. and C. Turchi. 2014. A comparison of supercritical carbon dioxide power cycle configurations with an emphasis on CSP applications. *Energy Procedia* 49:1187–1196

Neises, T., and C. Turchi. 2019. Supercritical carbon dioxide power cycle design and configuration optimization to minimize levelized cost of energy of molten salt power towers operating at 650 °C. *Solar Energy* 181:27–36. doi:10.1016/j.solener.2019.01.078.

Parrott, J. E. 1978. Theoretical upper limit to the conversion efficiency of solar energy. *Solar Energy* 21:227–29. doi:10.1016/0038-092X(78)90025-7.

Polimeni, S., M. Binotti, L. Moretti, and G. Manzolini. 2018. Comparison of sodium and KCl-MgC_{l2} as heat transfer fluids in CSP solar tower with sCO_2 power cycles. *Solar Energy* 162:51024. doi:10.1016/j.solener.2018.01.046.

Reyes-Belmonte, M. A., A. Sebastián, and M. Romero. 2016. Optimization of a recompression supercritical carbon dioxide cycle for an innovative central receiver solar power plant. *Energy* 112:17–27. doi:10.1016/j.energy.2016.06.013.

Saleh, B., and A. A. Aly. 2017. Screening of organic working fluids for a combined Rankine-refrigeration cycle driven by renewable energy. *International Journal of Applied Engineering Research* 12 (20):9575–86.

Saleh, B. A. A. A., A. F. Alogla, A. M. Aljuaid, M. M. Alharthi, K. I. E. Ahmed, and Y. S. Hamed. 2019. Performance investigation of organic Rankine-vapor compression refrigeration integrated system activated by renewable energy. *Mechanics & Industry* 20:206. doi:10.1051/meca/2019023.

Sarver, T., A. Al-Quraghuli, and L. L. Kazmerski. 2013. A comprehensive review of the impact of dust on the use of solar energy: History, investigations, results, literature, and mitigation approaches. *Renewable and Sustainable Energy Reviews* 22:698–33. doi:10.1016/j.rser.2012.12.065.

Shaaban, S. 2016. Analysis of an integrated solar combined cycle with steam and organic Rankine cycles as bottoming cycles. *Energy Conversion and Management* 126:1003–12. doi:10.1016/j.enconman.2016.08.075.

Shukla, A. K., A. Sharma, M. Sharma, and G. Nandan. 2018. Thermodynamic investigation of solar energy-based triple combined power cycle. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. doi:10.1080/15567036.2018.1544995.

Singh, H., and R. S. Mishra. 2018a. Performance analysis of solar parabolic trough collectors driven combined supercritical CO2 and organic Rankine cycle. *Engineering Science and Technology, an International Journal* 21:451–44. doi:10.1016/j.jestch.2018.03.015.

Song, J., X. Li, X. Ren, and C. Gu. 2018. Performance analysis and parametric optimization of supercritical carbon dioxide (S-CO2) cycle with bottoming Organic Rankine Cycle (ORC). *Energy* 143:406–16. doi:10.1016/j.energy.2017.10.136.

Tian, Y., and C. Zhao. 2013. A review of solar collectors and thermal energy storage in solar thermal applications. *Applied Energy* 104:538–53. doi:10.1016/j.apenergy.2012.11.051.

Wang, X., E. K. Levy, C. Pan, C. Wang, E. Romero, A. Banarjee, C. Rubio-Maya, and L. Pan. 2019. Working fluid selection for organic Rankine cycle power generation using hot produced supercritical CO_2 from a geothermal reservoir. *Applied Thermal Engineering* 149:1287–304. doi:10.1016/j.applthermaleng.2018.12.112.

Wang, X., Q. Liu, J. Lei, W. Han, and H. Jin. 2018. Investigation of thermodynamic performances for two-stage recompression supercritical CO2 Brayton cycle with high temperature thermal energy storage system. *Energy Conversion Management* 165:477–87. doi:10.1016/j.enconman.2018.03.068.

Wei, L., Y. Zhang, Y. Mu, X. Yang, and X. Hu. 2014. Simulation and experimental research of a low-grade energy conversion system using organic Rankine cycles. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 36 (5):537–46. doi:10.1080/15567036.2010.544001.

Xu, X., X. Wang, P. Li, Y. Li, Q. Hao, and B. Xiao. 2018. Experimental test of properties of KCl–MgCl2 eutectic molten salt for heat transfer and thermal storage fluid in concentrated solar power systems. *Journal of Solar Energy Engineering* 140 (5):051011. doi:10.1115/1.4040065.

Yagli, H. 2020. Examining the receiver heat loss, parametric optimization and exergy analysis of a solar power tower (SPT) system. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*. doi:10.1080/15567036.2020.1748765.

Yang, J., Z. Yang, and Y. Duan. 2020. Part-load performance analysis and comparison of supercritical CO2 Brayton cycles. *Energy Conversion and Management* 214:112832. doi:10.1016/j.enconman.2020.112832.

Yu., W., Q. Gong, D. Gao, G. Wang, H. Su, and X. Li. 2020. Thermodynamic Analysis of Supercritical Carbon Dioxide Cycle for Internal Combustion Engine Waste Heat Recovery. *Processes* 8:216. doi:doi:10.3390/pr8020216

Zhai, H., Y. J. Dai, J. Y. Wu, and R. Z. Wang. 2009. Energy and exergy analyses on a novel hybrid solar heating, cooling and power generation system for remote areas. *Applied Energy* 86:1395–04. doi:10.1016/j.apenergy.2008.11.020.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Phase evolution and microstructure of $BaBi_2Nb_2O_9$ ferroelectric ceramics

Ankita Banwal, Renuka Bokolia *

Department of Applied Physics, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:
Received 12 August 2020
Received in revised form 13 September 2020
Accepted 17 September 2020
Available online xxxx

Keywords:
BBN
Bismuth layered structure ferroelectric
Calcination
Ceramic
Sintering
Microstructure

ABSTRACT

In the present study, the phase evolution of $BaBi_2Nb_2O_9$ (BBN) composition prepared by solid-state reaction method have been studied. The detailed study for the formation of pure perovskite phase has been performed at a wide calcination temperature in the range of 700 °C–950 °C followed by sintering the ceramic at 1050 °C. The presence of a secondary phase is identified along with the formation of the desired $BaBi_2Nb_2O_9$ phase and is found to decrease with increasing calcination temperature. High resolution (HRXRD) depicts the evolution of single-phase $BaBi_2Nb_2O_9$ ceramic with an orthorhombic unit cell associated with the *Fmmm* space group at 950 °C, while SEM microstructural studies revealed the formation of dense plate-like grains on sintering the ceramic at 1050 °C. The optimum calcination temperature and sintering temperature were successfully investigated without any impurity traces for the application of high-temperature piezoelectricity and sensors.
© 2020 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Technological Advancements in Materials Science and Manufacturing.

1. Introduction

In recent years, Barium Bismuth Niobium Oxide BBN ($BaBi_2Nb_2O_9$) has attracted a lot of attention due to its various application in the high-temperature piezoelectric materials and Non-Volatile Random-Access Memory (NVRAM) because of their high curie point and fatigue-free properties respectively [1–5]. However, lead-based ferroelectrics are well known because of their piezoelectric properties and sensor applications [6,7], but they are toxic and not eco-friendly due to the presence of lead. So there is a need for alternative lead-free composition such as BLSF materials. The BLSF materials are bismuth layered structure ferroelectric, which is a member of the Aurivillius family having the general formula $(Bi_2O_2)^{2+}(A_{m-1}B_mO_{3m+1})^{2-}$, where m is signifying octahedral layers in perovskites block $(A_{m-1}B_mO_{3m+1})^{2-}$, embedded by Bismuth Oxide layer $(Bi_2O_2)^{2+}$, A represents monovalent cation or sometimes di or trivalent cations or a mixture of them such as Sr^{2+}, Ba^{2+}, Ca^{2+}, Pb^{2+}, Na^{+}, K^{+}, Bi^{3+}, Ln^{3+}, etc. whereas B denotes tetravalent, pentavalent or hexavalent transitions ions such as Ti^{4+}, Nb^{5+}, Ta^{5+}, W^{6+} which is suitable for octahedral co-ordinations [1–3,6,8]. The ceramic $BaBi_2Nb_2O_9$ (BBN) is BLSF with m = 2 having Ba and Bi ions at A-site and Nb ion at the B-site of the perovskite block $(A_{m-1}B_mO_{3m+1})^{2-}$. Most of the layered ferroelectrics are normal ferroelectric like $SrBi_2Ta_2O_9$, $SrBi_2Nb_2O_9$, $Bi_4Ti_3O_{12}$ but $BaBi_2Nb_2O_9$ and $BaBi_2Ta_2O_9$ are relaxor ferroelectric, which are characterized by diffuse phase transition [3,9,10]. The relaxor possesses high fatigue endurance and low remnant polarization along with high dc conductivity higher than any perovskite material [11,12].

The reported literature on the preparation of polycrystalline $BaBi_2Nb_2O_9$ ceramics showed a wide variation in the calcination temperature for obtaining the single phase of the compound. From the previous study, it is noted that the calcination temperature for BBN lies between 600 °C and 1000 °C [11–13]. It is reported that variation in calcination temperature influences the densification and crystallinity of the ceramic composition [14–16]. Hence, it is of immense interest to understand how the calcination temperature affects the phase evolution of the $BaBi_2Nb_2O_9$ ceramic powders and its microstructure and crystallinity. In this paper, we have optimized calcination temperature for BBN from 700 °C to 950 °C using a high-temperature furnace and sintered thee ceramic at 1050 °C to obtain a dense microstructure.

2. Experimental details

2.1. Fabrication

The ceramic $BaBi_2Nb_2O_9$ is prepared by a conventional solid-state reaction method. The required powder $BaCO_3$ (Sigma Aldrich

* Corresponding author.
E-mail address: renukabokolia@dtu.ac.in (R. Bokolia).

https://doi.org/10.1016/j.matpr.2020.09.380
2214-7853/© 2020 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Technological Advancements in Materials Science and Manufacturing.

Please cite this article as: A. Banwal and R. Bokolia, Phase evolution and microstructure of $BaBi_2Nb_2O_9$ ferroelectric ceramics, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2020.09.380

99.99%), Bi_2O_3 (Sigma Aldrich 99.99%), and Nb_2O_5 (Sigma Aldrich 99.99%) were weighed and mixed in stoichiometric amount. The mixed powder was then grinded in a mortar pestle with ethanol for 5 h. The grinded powder was calcined for 3 h at different calcination temperatures (700 °C, 750 °C, 800 °C, 850 °C, 900 °C, 950 °C) in a small alumina crucible. The calcined powder which has the least unreactive traces, was rigorously mixed with a standard binder (5 wt% polyvinyl alcohol) and then pressed into pallets of thickness 1 mm and diameter 10 mm with the help of a manual hydraulic press with pressure 50 MPa and resting for 1 min. The pressed pellets then sintered in a furnace at 1050 °C with intermittent heat treatment at 500 °C for 1 h to remove the binder from the pellets. The equation of synthesis reaction is represented by:

$$BaCO_3 + Bi_2O_3 + Nb_2O_5 \rightarrow BaBi_2Nb_2O_9 + CO_2 \uparrow \quad (1)$$

2.2. Characterization

The phase analysis of prepared samples of $BaBi_2Nb_2O_9$ ceramic was accomplished by using the X-Ray Diffraction spectra. The XRD spectra were obtained by using Cu-Kα radiation (λ = 1.5406 Å) in the Bruker D-8 Advance X-ray diffractometer. All the patterns were obtained in the interval 2θ ranging from 10° to 80° at a speed of 2° per minute. The phase was confirmed by matching all the peaks of prepared ceramic with the standard peaks in JCPDF file number #00-040-0355 using analytical expert high score plus software. The microstructure and grain formation was examined on an unpolished ceramic pallet using the JEOL 6610 Scanning Electron Microscope (SEM).

3. Result and discussion

Fig. 1 shows the evolution of the phases that are formed upon the initial calcination heat treatment of the powders ($Bi_2O_3 + Nb_2O_5 + BaCO_3$) when calcined at wide calcination temperature, T_c = 700 °C, 750 °C, 800 °C, 850 °C, 900 °C and 950 °C followed by sintering at higher temperatures range (T_s = 1050 °C) for 3 h to form the single-phase $BaBi_2Nb_2O_9$ compound using X-ray diffraction.

Raw powders mixed in stoichiometric ratio to form the $BaBi_2Nb_2O_9$, when calcined at 700 °C, shows the presence of mixed-phase of Bi_2O_3, $BiNbO_4$, and intermediate phase of $Bi_5Nb_3O_{15}$ along with some initial crystallization of the $BaBi_2Nb_2O_9$ phase (indicated by ◆ symbols) as shown in Fig. 1. It is reported that the formation of various intermediate phases (e.g., $Bi_5Nb_3O_{15}$, $BiNbO_4$, etc.) takes place when there occurs a reaction between Bi_2O_3 and Nb_2O_5 precursors [12]. It is evident from the XRD that when raw powders are calcined at 700 °C, show a clear presence of the Bi_2O_3 phase in the majority as compared to the $BaBi_2Nb_2O_9$ intermediate phases. The desired phase did not form distinctly at this temperature. The intermediate phase of $Bi_5Nb_3O_{15}$ was produced at 700 °C and coexisted with the desired single phase of $BaBi_2Nb_2O_9$ till 900 °C. The peaks corresponding to the Bi_2O_3 phase are observed due to the presence of unreacted Bi_2O_3. It is interesting to note that the relative intensities of the peaks corresponding to unreacted Bi_2O_3 and intermediate phase $Bi_5Nb_3O_{15}$, $BiNbO_4$ tends to decrease with the increase in calcination temperature from 700 °C to 900 °C. It may be noted that no intermediate phase corresponding to $BiNbO_4$ is observed when the sample is calcined at temperatures ≥900 °C. The reaction mechanism for the phase evolution of $BaBi_2Nb_2O_9$ completes at a higher calcination temperature of 950 °C, followed by sintering the powders at 1050 °C (Fig. 2).

The secondary phase contribution was quantified by taking the ratio of integral intensities of all the reflection corresponding to the secondary phases to the integral intensities of all the reflections of the diffraction pattern. The amount of secondary phase is found to decrease with increasing calcination temperature, as shown in Fig. 3. A close examination of the XRD traces revealed that samples

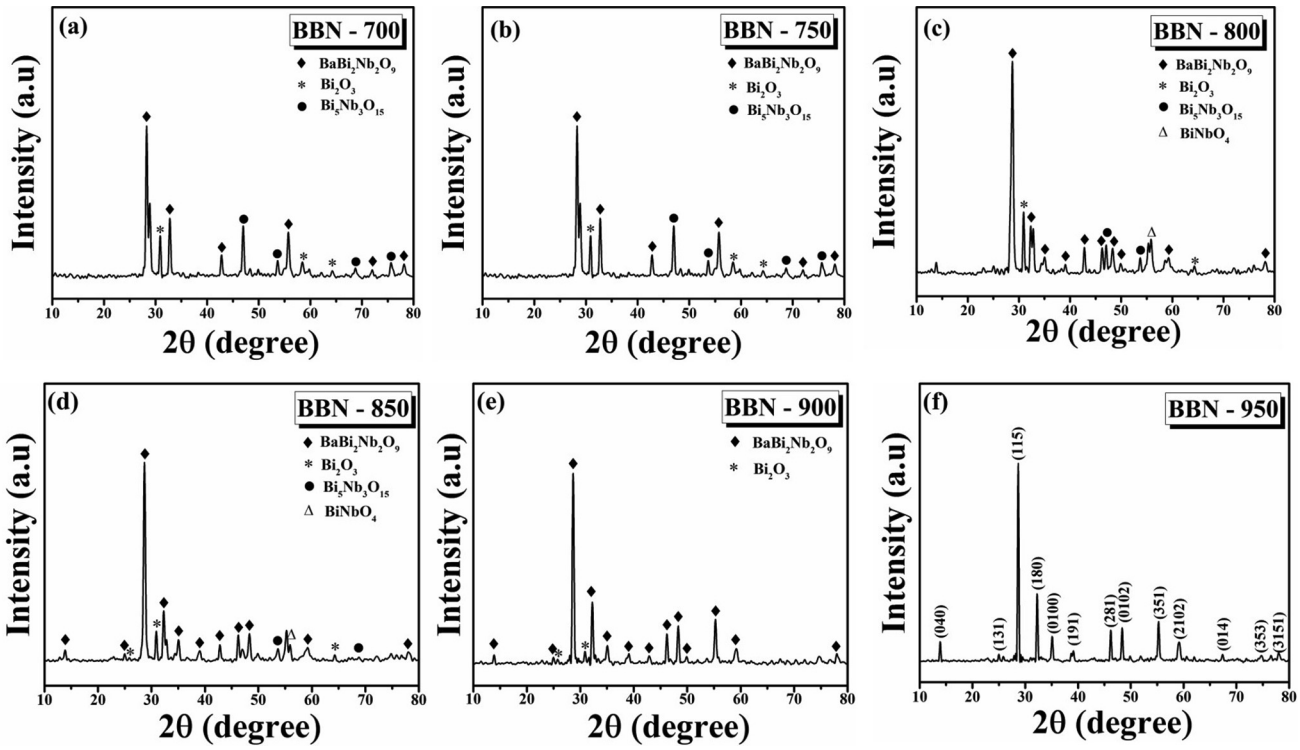


Fig. 1. (a-f) XRD patterns of $BaBi_2Nb_2O_9$ ceramic powders calcined at different calcination temperatures in the range of 700 °C–950 °C.

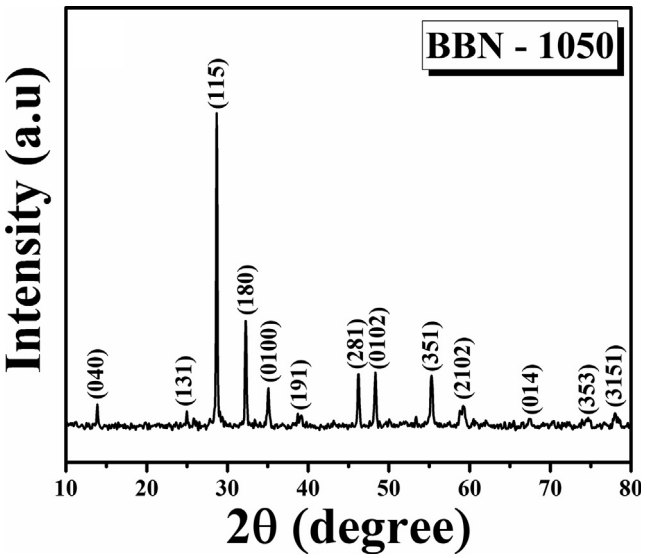


Fig. 2. Phase evolution of the sintered $BaBi_2Nb_2O_9$ ceramics at 1050 °C.

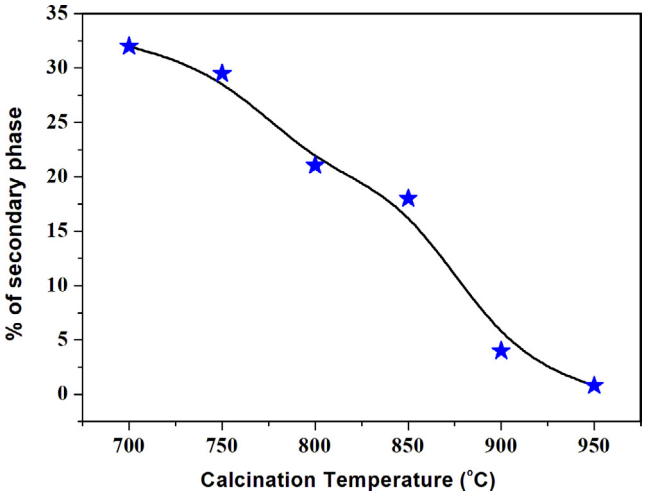


Fig. 3. Effect of calcination temperature on the formation of secondary phases.

calcined at temperature ≥900 °C exhibited all the XRD peaks pertaining to single-phase $BaBi_2Nb_2O_9$ and could be indexed and matched closely with the lone reference data on $BaBi_2Nb_2O_9$, in the JCPDS file of ICDD database (# 00-012-0403) as reported in earlier studies [2,3]. The pure phase of $BaBi_2Nb_2O_9$ ceramic with lattice parameters a = 5.5394 Å, b = 25.5751 Å, and c = 5.557 Å is associated with orthorhombic geometry with *Fmmm* space group [4,11]. Also, with an increase in calcination temperature, the full width at half maximum (FWHM) of the major Bragg's peak (1 1 5) becomes sharper, indicating the increase in the crystallite size and reveals the occurrence of grain growth, which is in agreement with the earlier reported results on $SrBi_2Nb_2O_9$ ceramics powders [12,17,18].

The microstructure of the sintered ceramic surfaces at 950 °C and 1050 °C observed from SEM images (Fig. 4) indicates the evolution of randomly oriented plate-like grains of varying size. However, this type of plate-like morphology has been commonly observed in the case of $BaBi_2Nb_2O_9$ ceramics owing to the anisotropic crystal structure in BLSF compounds [3,9]. To accentuate this fact further, the low surface energy along the (0 0 *l*) planes always tends to promote preferential growth of grains along the a–b plane, which leads to a plate-like morphology [6,19]. Thus, a higher calcination temperature of 950 °C seems to be the optimized calcination temperature, and sintering the ceramic composition at a higher temperature of 1050 °C results in improved crystallinity as well as dense microstructure of the ceramics.

4. Conclusion

The ceramic $BaBi_2Nb_2O_9$ prepared by solid-state reaction method investigated for detailed study of calcination temperature at different temperatures (700 °C, 750 °C, 800 °C, 850 °C, 900 °C, and 950 °C). XRD studies revealed the coexistence of different phases at different calcination temperatures. The presence of unreacted Bi_2O_3 along with intermediate phases $Bi_5Nb_3O_{15}$, $BiNbO_4$ were observed at lower calcination temperatures <900 °C. A close examination of the XRD traces revealed that samples calcined at temperature ≥950 °C exhibited all the XRD peaks pertaining to single-phase $BaBi_2Nb_2O_9$ with orthorhombic geometry. The surface morphology of the sintered ceramic at 1050 °C indicated the evolution of randomly oriented plate-like grains of varying size. Thus, a higher calcination temperature of 950 °C seems to be the optimized calcination temperature, and sintering the ceramic powders

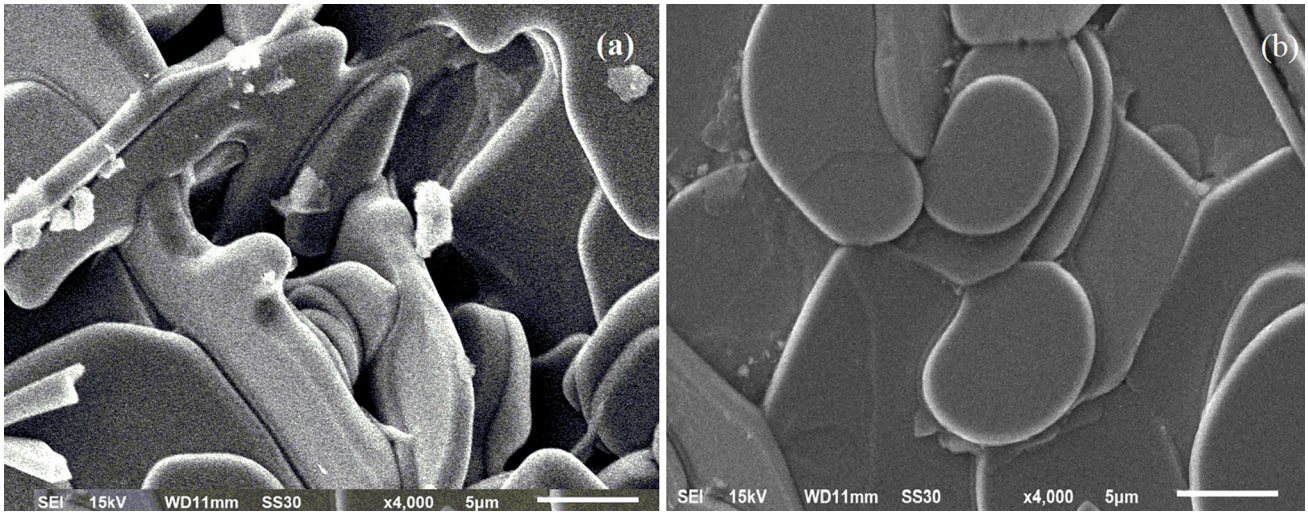


Fig. 4. SEM microstructure of sintered ceramic at (a) 950 °C (b) 1050 °C.

at a higher temperature of 1050 °C results in improved crystallinity as well as dense microstructure of the ceramics.

CRediT authorship contribution statement

Ankita Banwal: Conceptualization, Methodology, Investigation, Data curation, Writing - original draft. **Renuka Bokolia:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author would like to thank Delhi Technological University for providing all the research facilities and financial support under the DTU fellowship (2k18/PhdAP/505). Also, the author (AB) wishes to thank (RB) for all the guidance during this research work.

References

[1] R. Bokolia, O.P. Thakur, V.K. Rai, S.K. Sharma, K. Sreenivas, Ceram. Int. 41 (4) (2015) 6055–6066.
[2] M.X. Façanha et al., J. Lumin. 183 (2017) 102–107.
[3] B.R. Kannan, B.H. Venkataraman, Ceram. Int., 40 (PB) (2014) 16365–16369.
[4] Z. Zhang, H. Yan, X. Dong, Y. Wang, Mater. Res. Bull. 38 (2) (2003) 241–248.
[5] M.X. Façanha et al., J. Lumin. 204 (August) (2018) 539–547.
[6] A. Khokhar, M.L.V. Mahesh, A.R. James, P.K. Goyal, K. Sreenivas, J. Alloys Compd. 581 (2013) 150–159.
[7] T. Zheng, L. Luo, P. Du, A. Deng, W. Li, J. Eur. Ceram. Soc. 38 (2) (2018) 575–583.
[8] X. Hui, et al., Ceram. Int., 40 (8 PART A) (2014) 12477–12483.
[9] A.Z. Simões, B.D. Stojanović, Z.Ž. Lazarević, N.Ž. Romčević, J.A. Varela, E. Longo, Ferroelectrics 428 (1) (2012) 27–35.
[10] J.D. Bobić, M.M. Vijatović Petrović, J. Banys, B.D. Stojanović, Mater. Res. Bull. 47 (8) (2012) 1874–1880.
[11] M.P. Dasari, K. Sambasiva Rao, P. Murali Krishna, G. Gopala Krishna, Acta Phys. Pol. A 119 (3) (2011) 387–394.
[12] M. Verma, A. Tanwar, K. Sreenivas, J. Therm. Anal. Calorim. 135 (4) (2019) 2077–2087.
[13] D. Dhak, P. Dhak, T. Ghorai, S.K. Biswas, P. Pramanik, J. Mater. Sci. Mater. Electron. 19 (5) (2008) 448–456.
[14] R. Bokolia, O.P. Thakur, V.K. Rai, S.K. Sharma, K. Sreenivas, Ceram. Int. 42 (5) (2016) 5718–5730.
[15] D.E. López, K. Suwannakarn, D.A. Bruce, J.G. Goodwin, J. Catal. 247 (1) (2007) 43–50.
[16] T. Jardiel, A.C. Caballero, M. Villegas, Bol. la Soc. Esp. Ceram. y Vidr. 45 (3) (2006) 202–206.
[17] V. Garnier, I. Monot, G. Desgardin, Supercond. Sci. Technol. 13 (5) (2000) 602–611.
[18] W. Wu, S. Liang, X. Wang, J. Bi, P. Liu, L. Wu, J. Solid State Chem. 184 (1) (2011) 81–88.
[19] J.A. Horn, S.C. Zhang, U. Selvaraj, G.L. Messing, S. Trolier-McKinstry, J. Am. Ceram. Soc. 82 (4) (1999) 921–926.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Pre-Processing and Emoji Classification of WhatsApp Chats for Sentiment Analysis

Astha Mohta, Atishay Jain, Aditi Saluja, Sonika Dahiya
Computer Engineering Department
Delhi Technological University
Main Bawana Road, New Delhi, India
asthamohta@gmail.com, atishay9jain@gmail.com, saluja.aditi5@gmail.com, sonika.dahiya11@gmail.com

***Abstract*— WhatsApp is among the popular social media service with over 2 billion registered users. WhatsApp is integral in people's life as a medium of communication. They use WhatsApp to share their feelings through text messages. With WhatsApp present in over 180 countries, code-switching is common. Along with this, with the increase in usage and prevalence of emojis, emojis have become indispensable during sentiment analysis. Throughout this paper, our approach to convert unstructured WhatsApp messages to a structured form is discussed on which various data mining techniques for sentiment analysis can be performed. Our approach to deal with code-mixing, different emojis and the emotions they depict, and finally, perform basic analysis using this algorithm is discussed.**

Keywords— WhatsApp; Sentiment Analysis; Emoji; Interpersonal Communication; Emotion Analysis; Code-Mixing.

I. INTRODUCTION

WhatsApp is among the popular social media applications to connect people today, with 2 billion registered users across 180 countries and 55 billion messages sent per day [1], [2]. WhatsApp holds a key place in our life. People use WhatsApp to exchange information and also to share intimate feelings that are often touted as deep meaningful conversations.

In a study of students between the ages of 18 to 23 that was conducted in Chennai, India has investigated the importance of WhatsApp among youth. Students tend to spend 8 - 16 hours a day on WhatsApp [3]. Hence, can be concluded that WhatsApp chats are capable of depicting an individual's thought process as well as capturing their sentiments.

Pre-processing is an important step during the data mining of texts. Data pre-processing should be the first step for any process involving the use of data. That is because it leads to data sets, that are cleaner, coherent, and much more manageable, a must for any business trying to get valuable insights from the collected data.

Pre-processing data consists of various steps depending on the type of data. In the context of the work presented in the paper, the data obtained not only consists of textual but graphic data comprising of emojis as well. Hence, it becomes imperative that segregation is carried out before using them for any further analysis. Furthermore, the textual data obtained is multilingual and to enforce uniformity, translation to a common language is required. The above evidence leads to the requirement of a revised algorithm for pre-processing of these WhatsApp messages to use them in sentiment analysis. Conventional pre-processing for NLP does not account for bilingual speakers or code-switching. It does not acknowledge the importance of emojis and how they interact with the text either.

A key reason why WhatsApp is more popular than texts is that it is free. Along with this WhatsApp doesn't count words and supports multimedia [4]. Study shows that about 19.6% of tweets contain emojis and 37.6% users use emojis. Along with this, in 2015, "Face with Tears of Joy" [5] was the Word of the Year by Oxford dictionary.

Throughout the paper, our proposed algorithm to convert this unstructured data to a more structured form to perform various sentiment analysis algorithms is discussed. Emojis area categorized using both sentiment analyses in the following emotions: Happy, Sad, Angry, Fearful, Excited or Bored rather than which classify emotions as either positive, neutral or negative. For the study, people of different age groups were considered and their WhatsApp chats were analyzed (with their consent). The pre-processing on these chats are aimed to perform as well as count emojis in different categories to derive features from them to guide sentiment analysis.

In this paper, related work on pre-processing of data is first introduced used for sentimental analysis as well as the importance of emojis during texting and emotion categories for sentiment analysis in Section 2. Then presented our Data Set in Section 3. Our algorithm for pre-processing is discussed in section 4. Section 5 discusses our results and analyses on a sample of subjects. Conclusion and future work is presented in Section 6.

II. RELATED WORKS

WhatsApp came into the market as a substitute for SMS. It is used to make voice or video calls as well as to share media. WhatsApp provides these services for free [1]. Authors of "Survey Analysis on the Usage and Impact of WhatsApp Messenger" in their study of a group of WhatsApp users of ages between 18-50 found that about seventy-nine percent of their subjects use WhatsApp every day for at least 15-60 minutes [6]. In the study "Impact of WhatsApp on youth: A Sociological Study" 100 WhatsApp users were selected randomly, their ages lying between 18 and 30. The study

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

discovered that the frequency of usage for sixty-three percent users is fifty times a day, for twenty-one users are twenty times a day and sixteen percent is more than a hundred times a day. Thus can be concluded that amid the youth, WhatsApp has a huge influence on the way they communicate [7].

A study performed on the students of Abu Dhabi, eighty-five percent of women and seventy percent of men expressed that they used emojis as a substitute for facial expression on the text. Their average usage was 1 to 7 hours/day [8]. On a study on four million users of different countries, it was shown that about six million i.e., seven percent of all messages contained at least one emoji. This is a sign of the popularity of emoji among users [9].

In early research in 1982, Aravind K. Joshi describes code-switching or code-mixing as "Speakers of certain bilingual communities systematically produce utterances in which they switch from one language to another in the course of an utterance. Production and comprehension of utterances with Intra sentential code-switching is part of the linguistic competence of the speakers and hearers of these communities." [10]. "Code Mixing: A Challenge for Language Identification in the Language of Social Media" describes that despite most social media content being in the English Language, this still makes up for only half of the content worldwide. The reason for this is that most users switch between languages mid-way [11]. This leads to the need for a kind of translation or a model for tagging multiple languages.

For our study, six emotions have been chosen and classified our emojis in these categories. When looking at Russel's Circumplex Approach, he states that emotions can be classified on two scales, Valence and Arousal [12]. In this model, the emotions are represented as a combination of two dimensions: Valence, which represents a degree of pleasure - displeasure & Arousal, which represents levels of activation - deactivation. For example Boredom as an emotional state with negative valence, i.e. displeasure and low arousal [13].

For the study, six emotions have been chosen to convert all states of valence and arousal. As seen in figure 1, happiness and excitement can be seen in the quadrant with positive arousal (or pleasure) and valence (or activation), sadness and boredom have negative arousal (or displeasure) as well as low valence(deactivation) and finally, anger and fear have high arousal (or activation) but low valence (or displeasure). Low arousal with positive valence leads to a neutral state of mind [12].

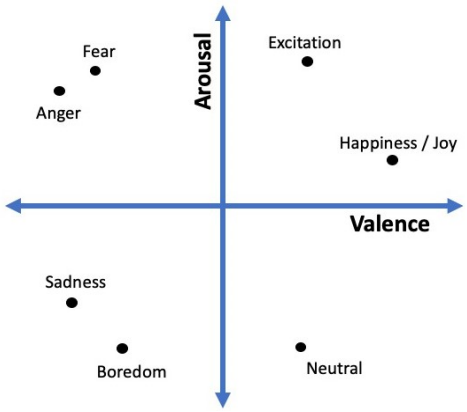


Fig. 1. Categorization of emotion based on Russell's Circumplex model

Authors of "Facebook Sentiment: Reactions and Emojis", argue that linguistic textual messages and emojis modify one another's meaning. The interaction of linguistic text with emojis can vary. An emoji can be a replacement for a phrase, reinforcement of a phrase, autonomously convey the emotion or attitude of the subject, provide emphasis on an already existing emotion in the message or can be used out of politeness. Given the context, emojis can be used to detect the subject's emotions [14]. The basic steps in any data pre-processing involve importing the libraries, data read, checking for unmitigated information, checking for categorical data, standardizing the data, PCA transformation & data splitting.

In the model prescribed in the paper "Social emotion mining techniques for Facebook posts reaction prediction", initial pre-processing like converting to lower case, removing hashtags etc. has been done using the Stanford CoreNLP Parser which is followed by the use of a tokenizer to split posts based on spaces & after filtering the stop words [15], the list of different tokens is obtained.

In the paper "Are Emoticons Good Enough to Train Emotion Classifiers of Arabic Tweets? ", showcases how emojis can be used to perform automatic labeling of tweets. Automatic labelling of tweets outperforms manual labeling of tweets [16].

III. DATA SET

A. Data Collection and User Consent

Eighteen individuals of various qualifications and ages were considered for analysis of their WhatsApp messages for the study. WhatsApp chats are private to everyone and hence certain ethical obligations are required to ensure their privacy and fair use. Each of them was asked to sign consent forms to permit the researchers to use their conversations for analysis. Their conversations are used for research purposes only no conversation has been made public. The complete data set is formed by collecting 5 – 10 chats from each participant. To collect the data, the "Export Chat" feature is used from WhatsApp that allows you to share your chats using Email, WhatsApp, Bluetooth, etc. [17]. The user has the option of sharing including media or not. Since worked only with textual data, data was shared without media.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

B. *Analysis of Subject*

Around 180 chats of 18 users are collected and stored in a database. For each participant, general information such as their names, age, gender and professional background is also collected. To present the diversity of the candidates selected for the study, analyzed them on their professional backgrounds and ages in Table 1 and Table 2.

TABLE I. NUMBER OF CANDIDATES BY AGE RANGE

Age Range	Number of Candidates
15 to 25	8
25 to 35	2
35 to 45	2
45 to 55	5
Greater than 55	1

TABLE II. NUMBER OF CANDIDATES BY PROFESSIONAL BACKGROUNDS

Professional Backgrounds	Number of Candidates
Engineering	5
Medicine	4
Humanities	4
Commerce	5

IV. PROPOSED ALGORITHM

The steps of the algorithm along with a sample chat are shown in figure 2.

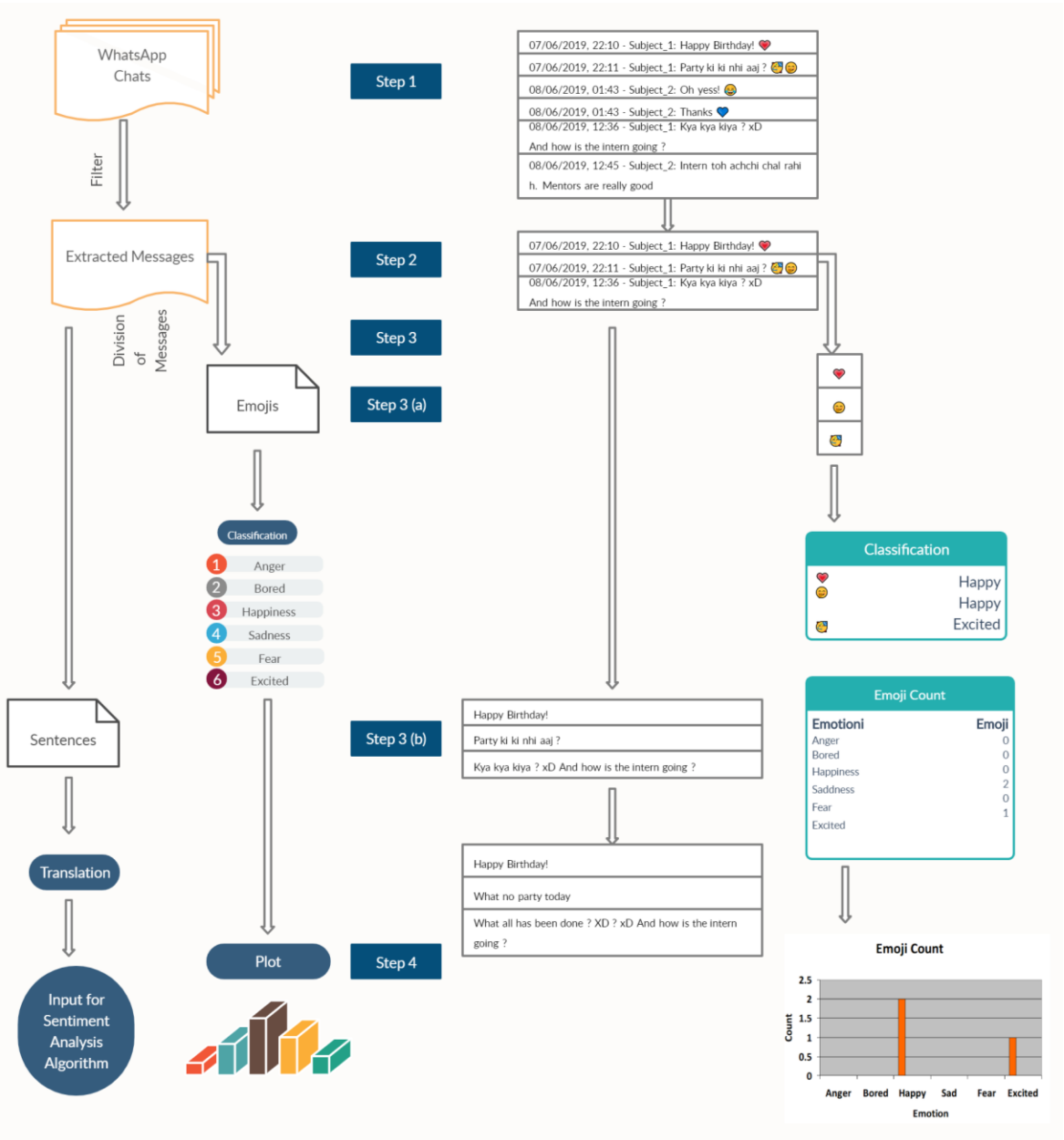


Fig. 2. Algorithm for Pre-Processing of WhatsApp Chats with example

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

The first step is data collection. After WhatsApp chats are taken from an individual, in the second step is to extract the message from the text file. To do so, first, the time stamp and date of the message is removed. Since a chat is between two individuals, but are only interested in analysis the sender side (our subject), the set is split into two parts using our subject name and pick the set, with messages from our subject.

For the next step, the messages are divided into two parts, one is the emoji and other is the sentence. The entire set of Unicode Emoji codes is supported along with aliases. An automatic labelling approach is used. Unicode emoji list has been used along with the keywords associated to classify them in six categories of emojis [18].

RegEx(Regular Expression) is a sequence of characters that represents a search pattern. RegEx can be used to check if a string contains the specified search pattern. In the work presented in the paper, the RegEx library in Python is used along with the Unicode library to figure out emoji patterns, distinguish between different emojis and classify emojis of the same pattern such as ones having the same structure but different colours as same [19]. Following is the list of Emojis classified:

Emotion	Emojis
Happiness	
Sadness	
Fear	
Anger	
Excited	
Bored	

Fig. 3. Emojis categorized by emotions

Emojis portray emotions unambiguously. Hence an emoji indicates emotion clearly and the emotion is both domain and topic independent. [20] Quite often though emojis are used relentlessly in a single text message. For this purpose, a threshold value was chosen after much deliberation and experimentation with various values. The threshold value chosen was 3. This implies that if the same emoji is used > 3 times after one another, it might lose its weightage. To take this factor into account, if emojis of the same classification is used > 3 times, the value is divided by 3 and add the ceiling value.

Let Emoji Count$_i$ = EC$_i$,

for i= [Bored, Anger, Sadness, Happiness, Excited, Fear,]

$$\text{then, } f(EC_i) = \{\ , \ = \begin{cases} \lceil EC_i/3 \rceil, & EC_i > 3 \\ EC_i, & otherwise \end{cases} \quad (1)$$

Emoji count for each emotion is initialized with zero. Using the classification in figure 3 the respective counts of categories as they are encountered is increased. These emojis can be depicted in a graphical form for pictorial analysis.

Following this, the set of sentences uses GoogleTrans (python library that implemented Google Translate API). GoogleTrans is a python library using which calls to such methods as detecting and translate is made using the Google Translate Ajax API to translate the input data to a particular language. Since the data collected in the work presented in the paper is multilingual, GoogleTrans API has been used to convert all of the text to a common language (English) before analyzing the data [21]. The number of words and sentences can be counted to give a rough idea about how interactive the subject is on WhatsApp. Along with this, the ratio of Sentence to Emoji is calculated to showcase if a subject uses emojis to better express themselves. These sentences can now be used as the input for and sentiment classification model.

V. Result Analysis

The algorithm proposed in the last section is applied to ten chats for each of the eighteen candidates. A sample of the comparison is presented below.

Two test subjects are analyzed and compare them based on the output of the pre-processed data approach discussed in the previous section. Each chat is analyzed by considering the total number of words, sentences & emojis present. Emoji Count in an online conversation is inversely related to the degree of professionalism between the participants [22]. They are often used as a replacement for words when felt and can't express our emotions with simple words, and are instead more comfortable using emojis [23].

In a study of how often introverts and extrovert tweet (in a fixed duration of time) it was discovered that the introverts tweet approximately 14.4% more than extroverts [24]. It is also studied that introverts prefer texting as a mode of communication over the phone calls, hence they text more. Consider a comparison between two Subjects from the dataset based on emoji count, word count, sentence count and the number of emojis used per sentence.

Table 3 gives a comparison based on words, sentences, emojis (along with categories) and emoji to sentence ratio.

TABLE III. Comparison of Subject 1 and Subject 2

Comparison	Subject1	Subject2
Sentence Analysis		
Word Count	142345	151969
Sentence Count	32647	27441
Emoji: Sentence Count	0.25	0.32
Emoji Analysis		
Total Emoji Count	8135	8819
Anger Emoji	340	259

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

Bored Emoji	496	571
Happy Emoji	343	762
Sad Emoji	176	363
Fear Emoji	6159	6214
Excited Emoji	621	650

It can be observed that Subject 2 has a greater emoji count per sentence compared to subject 1. Hence according to the studies mentioned above, it can be concluded that subject 2 mostly communicates unprofessionally and uses emojis as a method to express them better. This matches with subject 2's own opinion who stated that they use more emojis as they feel that emojis are more fun. The sentence count also reflects that texting is a mode of communication preferred by Subject 1. It has been observed that texting is more popular among introverts compared to extroverts. In the survey filled, subject 1 identified as an introvert, whereas subject 2 identified as an extrovert.

Similarly, when applied to all of the eighteen candidates, the result was compared with if the opinion of the subject themselves to see if it matched. The results are demonstrated in table 4.

TABLE IV. ACCURACY OF HYPOTHESIS

Correctly determined if Hypothesis matched with a subject opinion on if they are introvert/extrovert	**Count**	**Accuracy %**
Matched	14	77.78%
Did not Match	4	22.22%

VI. CONCLUSION

Data pre-processing is an important step in any machine learning algorithm. Converting raw data or unstructured data to a structured format is crucial before any further processing. Code-switching, which is quite prevalent in textual data obtained from social media poses a difficulty in processing, hence needs to be removed. The current methods do not account for this code-switching or how emojis interact with the text messages.

The algorithm explained in this paper successfully takes various WhatsApp chats as input and separates the texts of the two participants. This is followed by separation the emojis and sentences, & translating the sentences to a given language to make it unified and generic. The emojis are further classified for sentiment analysis using the formula proposed.

The results of the pre-processing have been used to analyze the social behavior of the subjects in context & finally classifying them as introverts or extroverts.

The data pre-processing technique described in this paper can be used for the sentiment & behavioral analysis. Blending a scientific approach for behavioral analysis with practical engineering goals of analyzing emotions in NLP texts can inspire a better and more intelligent approach to designing systems that interact with humans such as chatbots. Along with this psychologists can combine these approaches to enhance their models for sociological studies as well as in mental health care. In our next step, to design such a classifier is aimed, which gives a behavioral report of a person's emotions based on the emojis as well as a text sent by them.

REFERENCES

[1] J. Andjelic. (2019) Retrieved from https://fortunly.com/statistics/whatsapp-statistics#gref

[2] WhatsApp. Retrieved from https://www.whatsapp.com/about/

[3] Ka, Jisha & Jebakumar, Dr. (2014). Whatsapp: A Trend Setter in Mobile Communication among Chennai Youth. IOSR Journal of Humanities and Social Science. 19. 01-06. 10.9790/0837-19970106.

[4] Pinchas. (2018) Retrieved from https://www.telemessage.com/why-is-whatsapp-more-popular-than-sms-in-europe-infographic/

[5] Ljubešić, Nikola & Fiser, Darja. (2016). A Global Analysis of Emoji Usage. 82-89. 10.18653/v1/W16-2610.

[6] Kumar, Naveen & Sharma, Sudhahsh. (2017). Survey Analysis of the usage and Impact of Whatsapp Messenger. Global Journal of Enterprise Information System. 8. 52. 10.18311/gjeis/2016/15741.

[7] Bhatt, Anshu & Arshad, Mohd. (2016). Impact of WhatsApp on youth: A Sociological Study. IRA-International Journal of Management & Social Sciences (ISSN 2455-2267). 4. 376. 10.21013/jmss.v4.n2.p7.

[8] Kootbodien, Ammaarah & Prasad, Nunna & Ali, Muhamad. (2018). Trends and Impact of WhatsApp as a Mode of Communication among Abu Dhabi Students. Media Watch. 9. 10.15655/mw/2018/v9i2/49380.

[9] Lu, Xuan & Ai, Wei & Liu, Xuanzhe & Li, Qian & Wang, Ning & Huang, Gang & Mei, Qiaozhu. (2016). Learning from the ubiquitous language: an empirical analysis of emoji usage of smartphone users. 770-780. 10.1145/2971648.2971724.

[10] Aravind K. Joshi. (1982) Processing of sentences with intra-sentential code-switching. In Proceedings of the 9th conference on Computational linguistics - Volume 1 (COLING '82). Academia Praha, CZE, 145–150. DOI:https://doi.org/10.3115/991813.991836.

[11] Barman, Utsab & Das, Amitava & Wagner, Joachim & Foster, Jennifer. (2014). Code Mixing: A Challenge for Language Identification in the Language of Social Media. 10.13140/2.1.3385.6967.

[12] Acampora, Giovanni & Loia, Vincenzo & Vitiello, Autilia. (2011). A cognitive multi-agent system for emotion-aware ambient intelligence. IEEE SSCI 2011 - Symposium Series on Computational Intelligence - IA 2011: 2011 IEEE Symposium on Intelligent Agents. 0.1109/IA.2011.5953606.

[13] Posner J, Russell JA, Peterson BS. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev Psychopathol. 2005;17(3):715–734. doi:10.1017/S0954579405050340.

[14] Tian, Ye & Galery, Thiago & Dulcinati, Giulio & Molimpakis, Emilia & Sun, Chao. (2017). Facebook sentiment: Reactions and Emojis. 11-16. 10.18653/v1/W17-1102.

[15] Krebs, Florian & Lubascher, Bruno & Moers, Tobias & Schaap, Pieter & Spanakis, Gerasimos. (2017). Social Emotion Mining Techniques for Facebook Posts Reaction Prediction. 10.5220/0006656002110220.

[16] Hussien, Wegdan & Tashtoush, Yahya & Al-Ayyoub, Mahmoud & Al-Kabi, Mohammed. (2016). Are Emoticons Good Enough to Train Emotion Classifiers of Arabic Tweets?. 10.1109/CSIT.2016.7549459.

[17] WhatsApp. Retrieved from https://faq.whatsapp.com/en/android/23756533/

[18] Unicode (2020) Retrieved from https://unicode.org/emoji/charts/emoji-list.html

[19] Regex. Retrieved from https://regexr.com/

[20] Novak, P.K.; Smailovi´c, J.; Sluban, B.; Mozeti˘c, I. Sentiment of Emojis. PLoS ONE 2015, 10, e144296.

[21] SuHun Han. 2018. Retrieved from https://pypi.org/project/googletrans/

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

[22] Haji, Hadi & Nawxosh, Salam. (2019). The Use of Emoticons among University Students: A Pragmatic Study. Zanco Journal of Humanity Sciences. 23. 10.21271/zjhs.23.1.19.

[23] Alizah K. Lowell. (2016) Retrieved from https://www.psychologytoday.com/us/blog/contemporary-psychoanalysis-in-action/201605/why-do-we-use-emojis

[24] Zhou, Z., Xu, K. & Zhao, J. (2018) Extroverts tweet differently from introverts in Weibo. EPJ Data Sci. 7, 18 https://doi.org/10.1140/epjds/s13688-018-0146-8.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 06:13:49 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/345325553

Recognising Devanagari Script by Deep Structure Learning of Image Quadrants

Article *in* DESIDOC Journal of Library & Information Technology · November 2020

DOI: 10.14429/djlit.40.5.16336

CITATIONS
0

READS
27

2 authors, including:

Seba Susan
Delhi Technological University
75 PUBLICATIONS **376** CITATIONS

SEE PROFILE

All content following this page was uploaded by Seba Susan on 05 November 2020.

The user has requested enhancement of the downloaded file.

DESIDOC Journal of Library & Information Technology, Vol. 40, No. 5, Sept 2020, pp. 268-271, DOI : 10.14429/djlit.40.5.16336
© 2020, DESIDOC

Recognising Devanagari Script by Deep Structure Learning of Image Quadrants

Seba Susan* and Jatin Malhotra

Department of Information Technology, Delhi Technological University, Delhi – 110 042, India
**E-mail: seba_406@yahoo.in*

ABSTRACT

Ancient Indic languages were written in the Devanagari script from which most of the modern-day Indic writing systems have evolved. The digitisation of ancient Devanagari manuscripts, now archived in national museums, is a part of the language documentation and digital archiving initiative of the Government of India. The challenge in digitizing these handwritten scripts is the lack of adequate datasets for training machine learning models. In our work, we focus on the Devanagari script that has 46 categories of characters that makes training a difficult task, especially when the number of samples are few. We propose deep structure learning of image quadrants, based on learning the hidden state activations derived from convolutional neural networks that are trained separately on five image quadrants. The second phase of our learning module comprises of a deep neural network that learns the hidden state activations of the five convolutional neural networks, fused by concatenation. The experiments prove that the proposed deep structure learning outperforms the state of the art.

Keywords: Digitisation; Devanagari script; Handwritten character recognition; Structure learning; Part learning; Deep neural network; Convolutional neural network

1. INTRODUCTION

Digitisation of ancient Indian scripts is a part of the language documentation and digital archiving initiative of the Government of India. Ancient Indian languages were written in the Devanagari script that was in use in the 7th century C.E. It is considered as the origin of writing systems of several modern-day languages of the Indian sub-continent, such as Nepali, Hindi and Marathi[1]. Digitisation of handwritten scripts found in photographs is complicated since it is a multi-stage process involving several AI components. The first task is segmenting the text area from the image which is a challenging task by itself due to the presence of noise and artifacts such as the presence of foreign objects in the photograph[2]. The next step is to isolate individual letters from the text area by applying image processing techniques such as line detection and vertical and horizontal projection profiles as in[3] that segmented Brahmi letters from a script. After isolating individual characters, the next step is to employ an AI model that would recognise the character as resembling one of the alphabets of the language. This last and crucial step to digitisation and curation of documents, is well researched in literature, with several good surveys available on the topic[4-5]. In late 1980s, the neural network was used for learning directly from handwritten numeral images that were linearised before being given as input to the neural network[6]. From early 2000s, the focus was more on feature engineering. Specialised features for digit recognition were developed such as the box-partitioning method that divides the image into grids or windows, and computes measurements based on the area of the pixels occupied by the digit and the gradient of the image inside the window[7]. In yet another example, the mean normalised distance of each digit pixel from the bottom right corner of the window is computed[8]. In both these cases, the feature vector was formed by concatenating the values from all image sub-boxes, and learnt by a suitable classifier. In later years, a variety of popular image features such as HOG, SIFT, SURF, ORB and LBP were applied for the recognition of handwritten digits or characters[9-10].

In recent years, especially since 2012, deep learning[11] has become the trend in computer vision, with the focus being on 2D image processing using convolutional neural networks (CNN)[12], and increasing the number of layers to form deeper networks[13]. Constructing deeper and wider neural networks has led to an improvement in the classification of handwritten text as well, as seen from examples of Saha and Jaiswal who performed segmentation of Devanagari characters based on contours followed by classification with CNN[14], fine-tuning of pre-trained deep convolutional neural networks[15], capsule networks[16], deep belief networks[17] and deep convolutional neural network by Acharya *et al.*[18]. The deep CNN in[20] contained three repetitive layers of the combination of convolutional layer-rectified linear unit layer-normalisation layer-pooling layer. The problem with deep neural networks is that it lacks interpretability[21] and it is difficult to dissect the decision made by the deep CNN based on image content. Some possible solutions to the interpretability issue are structure learning[22] and part-wise semantic learning of the image[23]. We incorporate both structure learning and part-based learning in our work based on deep neural networks.

Received : 21 September 2020, Revised : 25 September 2020
Accepted : 29 September 2020, Online published : 4 November 2020

We focus on learning the Devanagari handwritten letters that includes the 36 alphabets and 10 numerals, using a two-phase classification module. In the first phase, five image quadrants are isolated and trained on separate Convolutional Neural Networks (CNN). The features extracted from the last dense layer of the convolutional network, for all five image quadrants, are concatenated, and fed as linearised input to a deep neural network (DNN) classifier. The classification scores are proved to be higher than the state of the art. The organisation of this paper is as follows. Section 2 introduces the proposed method. Section 3 discusses about the experimentation and the results. Section 4 gives the final conclusions and the future scope of this work.

2. PROPOSED METHOD

This work extends the idea of part-based learning introduced in a recent work of the authors on recognition of handwritten numerals from the MNIST dataset containing 60,000 English numeral images for training[19]. The 28 x 28 image was divided into five quadrants (a quarter of the image) -top left, top right, bottom left, bottom right and center. Each image quadrant was first linearised into a 1D input vector and was trained on a 196-35-35-35-10 Multi-layer perceptron (MLP) neural network. The hidden state activations (35-dimensional) of the last hidden layer was extracted from each quadrant, and concatenated to form a 175-dimensional feature vector that was trained on a support vector machine classifier. Our work was successful for the ten classes of English numerals. However, for the larger number of classes found in any language script, we have to incorporate deep learning for better text recognition results. This work advances our previous work in the following ways:

- The MLP neural networks used to train the five quadrants are substituted by five CNNs. The configuration of a single CNN is shown in Table 1. This ensures 2D processing of the image quadrant rather than 1D processing as in[19]
- The decision module in the second phase of our experiment comprises of a deep neural network (DNN) that is of configuration 2560-1000-1000-1000-46. This implies that there are three hidden layers each having 1000 neurons. The input layer is of dimension 512 x 5=2560. The number of output neurons is 46 since the number of classes of Devanagari characters are 46
- Our present work thus advances the state of the art by incorporating not only 10 digits, but the entire alphabet and 10 numerals of an ancient Indic script often found in Indian historical documents. We achieve this by incorporating deep learning into the five-quadrant system in[19] which is constricted for fewer classes (10 classes of numerals only).

Table 1. CNN architecture used for training each image quadrant

Layer	Kernel size	Output channels	Output shape	Stride	Activation function
Convolutional (2D)	3x3	16	12x12	1	Sigmoid
Max pooling (2D)	2x2	16	6x6	1	--
Convolutional (2D)	3x3	32	4x4	1	Sigmoid
Max pooling (2D)	2x2	32	2x2	1	--
Flatten	--	--	128		--
Dense (feature extracted)	--	--	512		--
Output	--	--	46		--

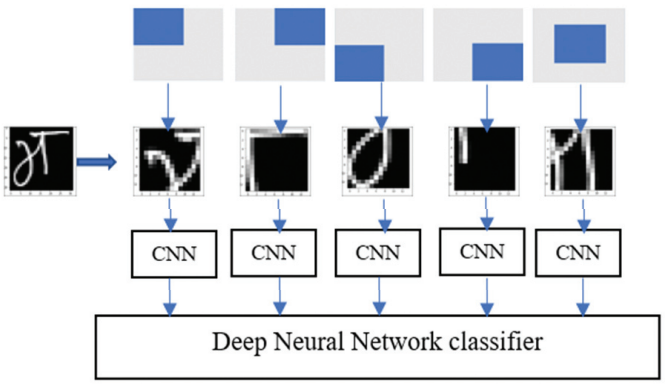


Figure 1. Proposed method.

The proposed approach is illustrated with the help of a block diagram in (Fig. 1). The figure shows the five quadrants -top left, top right, bottom left, bottom right and center. Each image is of size 28 x 28 and each image quadrant of size 14 x 14. Each 2D quadrant is trained on a CNN and the 512-dimensional feature vector (hidden structure) is extracted from the dense layer of each CNN, and concatenated to form the 2560-dimensional feature vector that is given as input to the DNN. The size of the DNN is 2560-1000-1000-1000-46.

3. EXPERIMENTATION AND RESULTS

The experiments are conducted in Python software on an Intel Core processor clocked at 2.8 GHz. The dataset is the UCI Devanagari handwritten character dataset[18] that has 46 categories of characters; samples from each of the 46 classes are shown in (Fig. 2). These include 36 alphabets and 10 numerals.

Overall, there are 92,000 images, with 2000 images for each of the 46 Devanagari characters. The dataset is segregated into 85 per cent training images and 15 per cent testing images at the source. This makes available 78,200 images for training and 13,800 images for testing. The CNN training and the subsequent feature extraction took 30 mins and the second learning module of deep neural network training took 15 mins for training on the entire dataset. The results are summarised in Table 2. As observed, the proposed network gives a high classification score of 98.39 per cent that is higher than the state of the art listed in Table 2. The only other method (Acharya *et al.* (2015))[18] that scores higher is trained on a larger training dataset, compiled by augmenting the data synthetically, by

Figure 2. Samples from 46 classes of Devanagari dataset[18].

Table 2. Performance comparison on the Devanagari test dataset

Method	Test accuracy (%)
Ours	98.39
Saha and Jaiswal (2020)[14]	93
Leathart *et al.* (2019)[20]	79.3
Acharya et al. (2015)[18] (*with five-fold increase of training data—Extended dataset*)	98.47
DenseNet pre-trained network[15]	90
VGG16 pre-trained network[15]	98
AlexNet pre-trained network[15]	98

five-fold. For the presently available training data of 78,200 images, our results show the highest accuracy.

4. CONCLUSIONS

A novel convolutional neural network based approach is proposed in this paper for recognizing Devanagari script from images. Each image is divided into five image quadrants, and each quadrant is trained on a CNN. The output of the dense layer in CNN is extracted and fed as input to a deep neural network. The results yield the maximum recognition scores for the Devanagari script. The contributions of our work are summarised as:

- The MLP neural networks used to train the five quadrants are substituted by five CNNs. This ensures 2D processing of the image quadrant rather than 1D processing as in[19]
- The decision module consists of the second phase of our experiment and comprises of a deep neural network (DNN) that is of configuration 2560-1000-1000-1000-46. This implies that there are three hidden layers each having 1000 neurons. The input layer is of dimension 512 x 5=2560. The number of output neurons is 46 since the number of classes of Devanagari characters are 46
- Our present work thus advances the state of the art by incorporating not only 10 digits but the entire alphabet and 10 numerals of an ancient Indic script often found in Indian historical documents. We achieve this by incorporating deep learning into the five-quadrant system in[19] which is constricted for fewer classes (10 classes of numerals only).

Application of our two-phase deep network module to other ancient Indic scripts such as Brahmi and Indus, forms the future scope of our work.

REFERENCES

1. Setlur, Srirangaraj. Guide to OCR for Indic Scripts. Edited by Venu Govindaraju. Springer London, 2009.
2. Susan, Seba & Devi, KM Rachna. "Text area segmentation from document images by novel adaptive thresholding and template matching using texture cues." *Pattern Anal. Appl.*, 2019, 1-13.
3. Singh, Ajay P. & Kushwaha, Ashwin Kumar. "Analysis of Segmentation Methods for Brahmi Script." *DESIDOC J. Libr. Inf. Technol.,* 2019, **39**(2).
4. Kumar, Munish; Jindal, Manish Kumar, Sharma, Rajendra Kumar & Jindal, Simpel Rani. "Character and numeral recognition for non-Indic and Indic scripts: A survey." *Artif. Intell. Rev.,* 2019, **52**(4), 2235-2261.
5. Jayadevan, R.; Kolhe, Satish R.; Patil, Pradeep M.; & Pal, Umapada. "Offline recognition of Devanagari script: A survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2011, **41**(6), 782-796.
6. Le Cun, Yann.; Jackel, Lionel D.; Boser, Brian; Denker, John S.; Graf, Henry P.; Guyon, Isabelle; Henderson, Don; Howard, Richard E. & Hubbard, William. "Handwritten digit recognition: Applications of neural network chips and automatic learning." IEEE Communications Magazine, 1989, **27**(11), 41-46.
7. Liu, Cheng-Lin.; Nakashima, Kazuki.; Sako, Hiroshi & Fujisawa, Hiromichi. "Handwritten digit recognition: Investigation of normalisation and feature extraction techniques.", *Pattern Recognition,* 2004, **37**(2), 265-279.
8. Susan, Seba & Singh, Veerendra. "On the discriminative power of different feature subsets for handwritten numeral recognition using the box-partitioning method." *In* 2011 Annual IEEE India Conference, pp. 1-5. IEEE, 2011.
9. Narang, Sonika Rani; Jindal, Manish Kumar & Sharma, Pooja. "Devanagari ancient character recognition using HOG and DCT features." *In* 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 215-220. IEEE, 2018.
10. Guruprasad, Prathima. "Handwritten Devanagari word recognition using robust invariant feature transforms." *In* 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 327-330. IEEE, 2015.
11. LeCun, Yann; Bengio, Yoshua & Hinton, Geoffrey. "Deep learning." nature, 2015, **521**(7553), 436-444.
12. Alex, Krizhevsky; Sutskever, Ilya & Hinton, Geoffrey E. "Imagenet classification with deep convolutional neural networks." *In* Advances in neural information processing systems, pp. 1097-1105. 2012.
13. Ian, Goodfellow; Bengio, Yoshua & Courville, Aaron. Deep learning. MIT press, 2016.
14. Saha, Pinaki & Jaiswal, Aashi. "Handwriting Recognition

Using Active Contour." *In* Artificial Intelligence and Evolutionary Computations in Engineering Systems, pp. 505-514. Springer, Singapore, 2020.
15. Aneja, Nagender & Aneja, Sandhya. "Transfer learning using CNN for handwritten Devanagari character recognition." *In* 2019 1st International Conference on Advances in Information Technology (ICAIT), pp. 293-296. IEEE, 2019.
16. Jeel, Sukhadiya; Suba, Yashi & D'silva, Mitchell. "Devanagari character classification using capsule network." *In* International Conference on Intelligent Systems Design and Applications, pp. 1040-1049. Springer, Cham, 2018.
17. Sazal, Md Musfiqur Rahman; Biswas, Sujan Kumar, Amin, Md Faijul & Murase, Kazuyuki. "Bangla handwritten character recognition using deep belief network." *In* 2013 International Conference on Electrical Information and Communication Technology (EICT), pp. 1-5. IEEE, 2014.
18. Acharya, Shailesh; Pant, Ashok Kumar & Gyawali, Prashnna Kumar. "Deep learning based large scale handwritten Devanagari character recognition." *In* 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pp. 1-6. IEEE, 2015.
19. Susan, Seba & Malhotra, Jatin. "Learning Interpretable hidden state structures for handwritten numeral recognition." *In* 2020 4th International Conference on Computational Intelligence and Networks (CINE), pp. 1-6. IEEE, 2020.
20. Tim, Leathart; Frank, Eibe; Pfahringer, Bernhard & Holmes, Geoffrey. "On calibration of nested dichotomies." *In* Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 69-80. Springer, Cham, 2019.
21. Zhang, Quan-shi & Zhu, Song-Chun. "Visual interpretability for deep learning: a survey." *Front. Inf. Technol. Electron. Eng.,* 2018, **19**(1), 27-39.
22. Hou, Bo-Jian & Zhou, Zhi-Hua. "Learning with interpretable structure from rnn." arXiv preprint arXiv:1810.10708 (2018).
23. Quanshi, Zhang; Cao, Ruiming; Wu, Ying Nian & Zhu, Song-Chun. "Growing interpretable part graphs on convnets via multi-shot learning." *In* Thirty-First AAAI Conference on Artificial Intelligence. 2017.

CONTRIBUTORS

Dr Seba Susan is an Associate Professor in the Department of Information Technology at Delhi Technological University (DTU). She completed her Ph.D from the Electrical Engineering Department of Indian Institute of Technology (IIT) Delhi in 2014. Her research areas are Machine Learning, Computer Vision, Speech and Natural Language Processing.

Mr Jatin Malhotra is a postgraduate student in Information Systems in the Department of Information Technology at Delhi Technological University (DTU). His areas of interest are Machine learning and Image processing. His current research involves developing deep learning modules for handwriting recognition.

View publication stats

Volume 8 Issue 2 (2020) 82-85 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Review on Sheet Metal Formability

Shivang Chhikara, Rajat Khatta, Suraj Prakash Verma, Dadge Mukesh Shamrao
Department of Mechanical Engineering, Delhi Technological University, Bawana Road, New Delhi-110042, India

Article Info

Article history:
Received 17 January 2020
Received in revised form
20 May 2020
Accepted 28 May 2020
Available online 15 June 2020

Keywords: Formability, Forming Limit Diagram, Yielding, Anisotropy, Strain Rate

Abstract: Sheet metal Forming is one of the most important and diverse manufacturing processes and has been an integral part of industries like automotive and aerospace which need high density, lightweight and high strength materials. This paper tries to encapsulate the features related to Formability and its testing. Various parameters like anisotropy and strain rates on which the formability depends are also considered and their impact on the testing. Further, the paper includes study of Formability of Tailor Welded Blanks which is an integral part of the automotive sector. At last the paper concludes with the future developments that can be made in this field along with new modelling techniques and new advancements in testing.

1. Introduction

Sheet-Metal Formability is generally defined as the ability of the material (typically called as blank) to undergo plastic deformation governed by flow rules to give the desired shape change without failure, such as wrinkling, cracking, necking, cracking, tearing, etc. Formability is not easily quantified as it depends on several factors such as material flow properties, ductility, die geometry, die material, lubrication, press speed, stress rate, strain rate, temperature roughness, springback, strain localization, etc. [1]. Further, formability of different materials also depends upon the sheet metal process that is being undertaken for the production of the part. Important processes in this field are bending, deep-drawing, stretch forming, hydroforming, incremental forming and many more. The formability of any sheet is at its minimum at plane strain conditions and so about 80-85% industrial failures occur around this state. So, an ideal test should be able to simulate these conditions efficiently and effectively. It should also be taken into consideration that no single formability test can relate the formability for all types of forming or stamping operations [2]. The basic forming characteristics of sheet metals can be obtained from simple mechanical tests such as tensile, bulge and hardness tests but these tests are too simple to relate to various parameters and phenomena such as strain-hardening. A few important parameters on which the formability of a sheet metal depends on are the strain-hardening index (n) which is usually obtained from flow curves (fig.1), strain rate and anisotropy of the material. Higher the strain hardening index and higher the strain-rate

strain before necking. Further, improvements in the drawability of the material is seen when the Anisotropy index (r) is higher. But there is a major disadvantage of these simple tests, i.e., they completely ignore the effect of variables (geometry of punch or die, lubrication, punch speed etc.) on the forming behavior of sheet metal completely. According to the power-law of strain hardening-

$$\sigma = k\varepsilon^n \qquad (1)$$

2. Formability Tests

All the sheet metal industries have been looking for an ideal formability t
splitting failure that usually occur in plane-strain condition [1]. The most accepted way to predict of as sheet metal is by plotting the Forming Limit Diagram (FLD) (fig.2). It is an experimental procedure in which a combination of two principal surface strains (major and minor) are represented. The diagram consists of Formability Limit Curve (FLC) above which localized instability is observed. The major and minor strains plotted in the FLC are obtained from these tests that are done under various train paths from uniaxial tension over plane strain tension to biaxial tension [4]. In stretching is measured and the metals are rated on the basis of them. Formability tests are mainly divided into three broad categories- intrinsic tests, simulative tests and the test that are used to plot the FLDs. The intrinsic tests consist of basically the uniaxial tension tests and are independent of the sheet metal geometry. In simulative tests, drawing and stretching conditions are simulated because these are the most common ones in the industrial applications. The different types of simulative tests are Erichsen Olsen Tests, Swift Cup test, Limit Dome Height test and OSU tests. There are two experimental methods for obtaining FLDs, one is, Out-of plane formability tests [5,6] and the second one is, In-plane formability tests [7,8]. For both the tests, a grid marking scheme is used in which the sheet metal is marked with a grid pattern and is deformed in the stretching tests [9]. After this the deformation of the grid pattern is measured where either necking has taken place or fracture has occurred, thus, giving the values of major and minor strains. Examples of these tests include Nakajima Test [10], Marciniak Test [11].

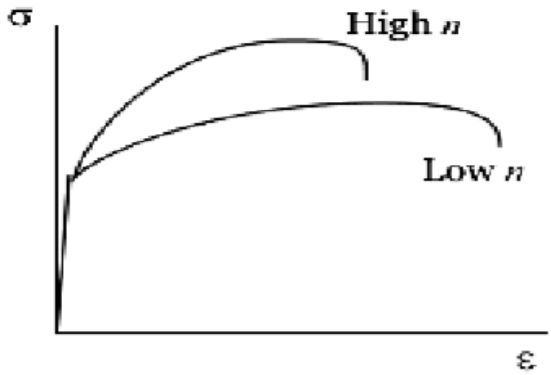


Fig.1: True stress versus True Strain Diagram [3]

Various theoretical models are also proposed to calculate forming limit diagrams of different sheet metals. These Modelling techniques are mostly based Bifurcation theory, geometrical imperfection theory and Continuum damage mechanics [4]. The modelling is also an integral part of formability analysis because during the process of sheet metal forming, the plastic deformation itself includes a lot of different phenomena like anisotropy, hardening, failure and fracture that too occurring simultaneously. Thus, modelling or numerical analysis tries to combine all these factors so that the cumulative effect on formability can be seen. Some of the important modelling
l (Maximum
force criterion), Modified maximum force criterion (MMFC), Storen and Rice (S-R) theory, Marciniak and Kuczynski (M-K) model and one of the recent developments in this field is NADDRG model.
Swift cup test evaluates the drawability of a sheet metal by simulating the drawing operation, in which the sheet is held under a blank holder and is drawn into small parallel sided cup. Drawability is measured by drawing various blanks with increasing the diameter one by one and thus defines the Limit Draw Ratio (LDR) which gives an estimation of the drawability. Swift cup test is highly standardized,

***Corresponding Author,**
E-mail address: shivangchhikara@gmail.com
All rights reserved: http://www.ijari.org

with die and punch radii as 6.35mm. But its drawback is that it is a very time-consuming process

$$LDR = \frac{Maximum\ Blank\ Diameter}{Cup\ Diameter} \quad (2)$$

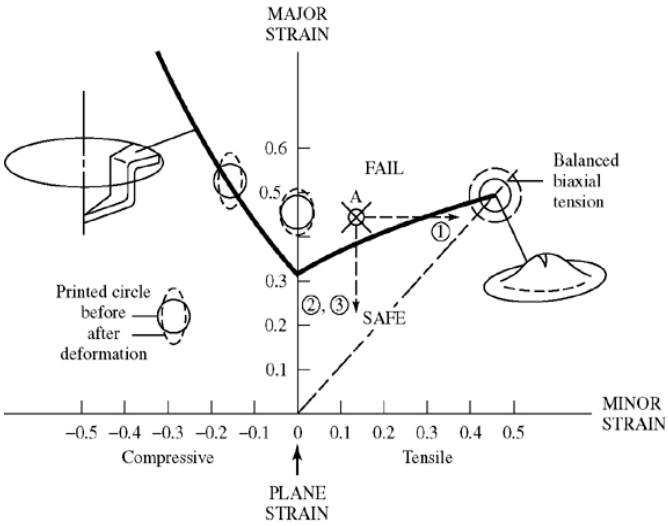


Fig.2: Forming Limit Diagram (FLD) [12]

Erichsen and Olsen are basically two tests that simulate the stretching operation conditions. In these tests, a sheet metal is clamped between two plates and a hole of diameter 25.4 mm is made into it. Now, a ball with diameter, d, is pressed into the sheet until the fracture occurs. For Olsen test, d=22.2 mm and for Erichsen test, d=20mm. Height of the cup, h at failure is used as formability index, higher this index is higher will be the formability of the material. The disadvantages of this test are first they cannot be correlated with other mechanical properties of the sheet metal and second these tests have poor data reproducibility. Limiting Dome Height Test (LDH) simulates the stamping conditions more effectively. This test includes clamping the metal sheets of varying width into a blank holder and a 102 m hemispherical punch is used to stretch the sheet over. Small diameter circles (2.5mm) are marked like a grid on the sheet and the circle closest to the fracture gives the width strain. The height at which dome fails gives us Limiting Dome Hight (LDH) near plane strain and thus is used as a formability index. The standard width used in the industry is 133mm. The disadvantages of this test are that it gives scattered results on various machines and tooling. Further, stable plane strain conditions are not achieved in this test due to the axisymmetric tooling. In OSU (Ohio State University) test there is no axisymmetric tooling of punch diameter of 25.4 mm thus giving stable plane strain conditions. In this test, 5 sheets are used of width 124mm and length 178mm and results are averaged. This test also defines formability index at the height at which failure occurs. OSU test is said to be 6 times faster than the LDH test.

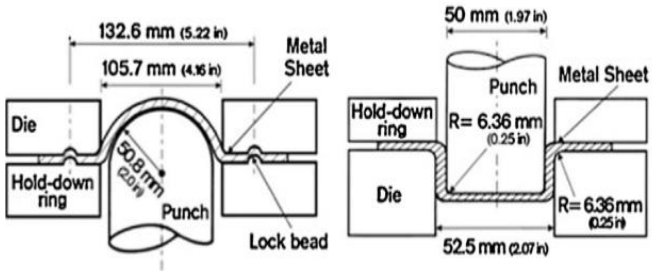


Fig.3: Sketch of the limit dome height test on the left [13], and the sketch of swift cup test on the right [14]

3. Anisotropy and Formability

Till the 1980s the yield criterion for materials did not have the required anisotropic coefficients that these criterions were not much accurate. Hills provided the useful yield criteria for anisotropic materials in 1948 and it was the only one to have some applicability in a number of stress cases.

The isotropic yield conditions for various metals were given by $_i$) and deviatoric stress (Sj) tensors.

$$\phi = |\sigma_1 - \sigma_2|^a + |\sigma_2 - \sigma_3|^a + |\sigma_3 - \sigma_1|^a = |S_1 - S_2|^a + |S_2 - S_3|^a + |S_3 - S_1|^a = 2\underline{\sigma}^2 \quad (3)$$

$_{1,\ 2,\ 3}$ are the principal stresses, $\underline{\sigma}$ is the effective stress.

For a=2 and a=4 the equation becomes the Von mises criterion for a=1 and for the limiting case of $a \to \infty$ the equation becomes

For isotropic materials the yield criterions are independent of the reference frames. But for anisotropic materials yield properties are directional and thus the yielding criterion is independent on the reference frame. Thus, Hill extended the Von Mises criteria [16] to orthotropic materials (which exhibit orthotropic symmetry) $\phi = F\left(\sigma_{yy} - \sigma_{zz}\right)^2 + G\left(\sigma_{zz} - \sigma_{xx}\right)^2 + H\left(\sigma_{xx} - \sigma_{yy}\right)^2 + 2\left(L\sigma^2{}_{yz} + M\sigma^2{}_{zx} + N\sigma^2{}_{xy}\right) = \underline{\sigma}^2$ (4)

Here, F, L, G, H, M, N are material-based constants. This yield criteria provided accounts for planar anisotropy which is obtained from the equation (5)

$$r_0 = \frac{r_0 + 2.r_{45} + r_{90}}{4} \quad (5)$$

These values of r are determined by uniaxial tension tests and then cutting the sheet metal along three directions in the plane of the sheet at 0˚, 45˚ and 90˚.

Lankford et. Al [17] introduced the term strain ratio, or as it is commonly called as the plastic anisotropy r-value and used this to represent the drawability of different sheet metals.it is also seen in some researches that the limiting draw ratio in simulative tests is quite closely related to the average r-value. Thus r-value of anisotropy has been quite extensively used as a parameter for the drawability of sheet metals [17].

4. Formability and Strain Rates

Hecker [18] proposed that strain-hardening index (eqn. 6) was not alone responsible for the prediction of fracture height in Olsen test and thus strain-rate sensitivity index also has to be taken into consideration. Thus, it was seen that for a given strain-hardening index (n), higher the strain-rate sensitivity index (m), the material will yield a larger cup height, therefore, indicating an increase in formability of the sheet metal [19]. Hence, for a given value of n steel has more formability than aluminum since it has higher strain-rate sensitivity index

$$m = \frac{\partial ln ln\, \sigma}{\partial \varepsilon} \quad (6)$$

It is seen that formability increases with high speed due to the reasons such as constitutive behavior, inertial effect and die impact effect. Also, some materials show increase in ductility and low stress too. Strain rate effects also play a role in increased formability due to high-velocity [20]. Due to inertial aid, there is a change observed in the necking region which causes the increase in formability of sheet metal [21]. Dariani et al. [20] conducted a study on FLDs of Al 6066-76 & AISI 1045 sheets at various deformation velocities int the explosive forming process and found that formability increased by 53.8% by impact and 146.1% by explosive forming. Gerdooei & Dariani [22] analyzed the formability of metal sheet at varying strains (from 0.001 to 100/s) and their effects were shown on the critical plane strain (FLD_0).

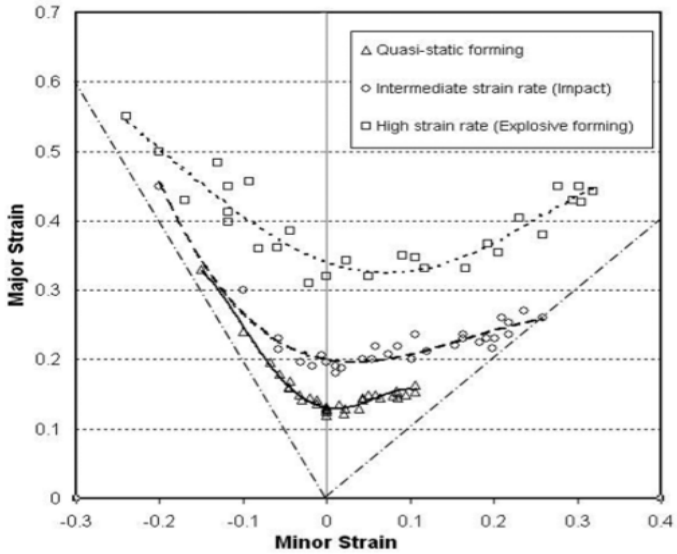


Fig. 4: FLDs of Al 6061-T6 sheet under various strain rates

5. Formability in Tailor Welded Blanks

As we are seeing, there is a worldwide trend of trying and achieving low weight, high density, high strength materials with high formability in industries such as automotive and electrical applications. Therefore, the use of Tailor Welded Blanks (TWBs) have increased tremendously as this process allows materials of varying thickness, properties or surface conditions in a single pressed or stamped part [23]. The parts produced by this process have joints of high strength and high reliability. In making Tailor Welded Blanks, 99.9% of the times either CO_2 or Nd: YAG laser welding is used [24]. But theses welding processes affect the formability of the sheet metal or the end product too. Mustafa A. Ahmetoglu et. Al [25] concluded in their study that in case of TWBs failure in deep drawing usually occurs at the flat bottom of the punch parallel to the weld line and this happens due to the non-uniform distribution of the deformation. According to Hisashi Kusuda et. al. [26], in automotive sector, TWBs are used in making three different parts- 1. Side panel, outer 2. Door, inner 3. Front side inner and the formability problems seen in these three respectively are split in stretch flanging, stress concentration in thin plate and insufficient ductility of weld bead. Thus, by considering these 2 studies, a point is very clear that there is a constant need of new formability data for the TWBs to control the failure of the formed parts. Saunders et. al. [27] conducted four different formability tests on the Tailor Welded Blanks by using two base materials, first, 1.8 mm Aluminum- Killed Drawing Quality (ASDQ) steel and second, 2.1 mm High Strength Low Alloy Steel (HSLA) and concluded that two modes of failures were occurring. Type 1 failure occurred when the principal stretch axis lied parallel to the weld and fracture depended on the ductility. Type 2 failure was the dominant forming failure which was not affected by welding operations. Their study further concluded that weld-line displacement is also an important parameter of formability in TWBs.

6. Future scope in formability analysis of sheet metal forming

At the present stage, different FLC models have different advantages to offer and at the same time their inherent limitations also. So, the future of FLCs would be to develop new models that can increase their advantages and at the same time suppressing their limitations. As seen in the discussion of strain rates it is evident that formability of metals improves with high velocities to a great extent. This has led to researchers focus on techniques of high velocity forming as seen in processes like Explosive forming, Electromagnetic Forming, etc. Some non-conventional testing methods are also emerging nowadays such as Continuous Bending under Tension (CBT) test which generates cyclic stretch bending [28]. Tube Extension tests are also been put to use for the study of formability of sheet metals. Also tube forming is nowadays used in automotive radiator closure also and the advantage lies in the fact that this process has fewer components than the conventional process and stiffness of the parts are also significantly increased. Hydroforming is also one of the new techniques used for tubular products and enables testing under bi-axial stress state. Further techniques such as Incremental Forming are also gaining momentum because this process uses universal tooling and thus increases the flexibility of the forming method. Further, materials at very high temperatures are also formed by various sheet metal processes and the formability calculated under such circumstances is of utmost importance. So, formability analysis of such new techniques is also needed to ensure that there is reduction in costs and low scrap is generated and the forming methods could be used to different set of materials. So, at present the main focus of researchers and engineers should be on developing a more sophisticated testing equipment so that a single test ensures procurement of maximum possible information of formability of the sheet metal. Also, new modelling techniques and numerical analyses should also be developed which strengthen the theoretical concept of FLDs and can confine a lot more forming parameters and their consequent effects. Further, advancements in the field of new materials should also be looked forward to, as the new materials open new horizons of high strength and lightweight application and thus can be beneficial to the industry. And this can further increase the profits of the organizations.

7. Conclusions

This paper tries to give an insight to the reader about the importance of sheet metal forming methods and how formability analysis is an integral aspect of it. The paper gives a brief overview of the various tests involved in getting quantifiable data on formability and to get the forming limits above which the failure occurs. There has been a lot of advancements in formability over the past four decades and still the research continues in search of new materials, comprehensive testing methods & equipment and mathematical analyses of the yielding criterion considering a lot more parameter and their effects on formability. So, this paper has been an attempt to study the basic nature of formability and its importance in the industry of today.

References

[1]. K. Narsimhan, M P Miles, R.H. Wagoner, A better sheet-formability test, 1995
[2]. R.M. Hobbs & J.L. Duncan in: P.D. Harvey (Ed.), A course on advanced sheet metal Forming, ASM Metals Park, OH, 1979
[3]. Sheet Metal Forming Technology, Dr. NS Mahesh
[4]. Ruiquiang Zhang, Zhutao Shao, Jianguo Lin, A review on modelling techniques for formability prediction of sheet metal forming
[5]. ISO 12004-2, Metallic material-sheet and strip- determination of Forming Limit Curves. Part 2: Determination of forming limit curves in laboratory, 2008, pp. 3-8
[6]. P.F. Bariani, S. Bruschi, A. Ghiotti, A. Tureta, Testing Formability in the hot stamping of HSS, CIRP, Ann. 57(2008) 265-268
[7]. Z. Marciniak, K. Kuczynski, Limit strains in the process of stretch-forming sheet metal, Int. J. Mech. Sci. 9 (1967) 609-620
[8]. K. Raghvan, A simple technique to generate in-plane forming limit curves and selected applications, Metall. Mater. Trans. 26 (1995) 2075-2084
[9]. Testing and modelling of material behavior and formability in sheet metal forming, S. Bruschi, T. Altan, D. Banabic, P.F. Bariani, A. Brosius, J. Cao, A. Ghiotti, M. Khraisheh, M. Merklein, A.E. Tekkaya
[10]. Kikuma T. Nakazima K (1971), Effects of the deformation conditions and mechanical properties on stretch forming of steel

sheets. Transactions of the Iron and Steel Institute of Japan 11: 827-839
[11]. Marciniak Z., Kuczynski K., Pokota T. (1973) influence of plastic properties of a material on the forming limit diagram for sheet metal in tension. International Journal of Mechanical Science 15: 789-805
[12]. M. El Sherbiny a

(2014) 869-879
[13]. Ayres R.A, Braier WG, Sajewski VF (1979) Evaluating the limit dome height as a new measure of press formability, Journal of Applied Metalworking 1(1): 41-49
[14]. Swift HW (1952) Plastic instability under plane stress, Journal of the Mechanics and Physics of Solids 1:1-18
[15]. Dorel Banabic, Frederic Barlat, Oana Cazacu, Toshihiko Kuwabara, Anisotropy & Formability
[16]. Hill, R: A theory of the yielding and plastic flow of Anisotropic Metals, Proc. Roy. Soc. London A193 (1948) 281-297
[17]. Y.G. An, H. Vegter, S. Melzer, P. Romano Triguero, Evolution of the plastic anisotropy with straining and its implication on formability for sheet metals
[18]. SS. Hecker, Metallurg. Eng. Quart., 14 (1974) 30
[19]. Hans Conrad, Effect of the strain-rate sensitivity of the flow stress on stretch formability of sheet metals.
[20]. B.M. Dariani, G.H. Liaghat, M Gerdooei, Experimental investigation of sheet metal formability under various strain rates
[21]. Hu. X and Daehn G.S., Effect of velocity on localization in uniaxial tension, Acta Mater., 1996, 44, 1996,1021
[22]. Gerdooei M and Dariani B.M., Dynamic analysis instability of sheet metal under biaxial stretching. Amirkabir J. Sci. Technol., 2007, 18 (67-B), 31-39
[23]. K. Schmidt, Growing sheet me stamping blanks, metal forming, (September 1994) ,43-46
[24]. E. Ahmed, U. Resgen, M. Schleser and O. Mokrov, on formability of Tailor laser welded blanks of DPI/RIP steel sheets
[25]. Mustafa A. Ahmetoglu, Dirk Brouwers, Leonid Shulkin Laurent Taupin, Gary L. Kinzel, Taylan Altan, Deep drawing of round cups from tailor welded blanks
[26]. Hisashi Kusuda, Toshiyuki Takasago, Fumiaki Natsumi, Formability of tailored welded blanks
[27]. F.I. Saunders & R.H. Wagoner, Forming of Tailor welded blanks
[28]. Emmens W, Van den Boogerd AH (2009) Cyclic stretch Bending: Mechanics stability and formability

2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)
Galgotias University, Greater Noida, UP, India. Oct 30-31, 2020

RNA Secondary Structure Prediction using Machine Learning: A Review

Shubham Mittal
Department of Biotechnology
Delhi Technological University
Delhi, India
mshubham9606@gmail.com

Yasha Hasija*
Department of Biotechnology
Delhi Technological University
Delhi, India
yashahasija06@gmail.com
*corresponding author

***Abstract*—Complete knowledge of base pairing in RNA secondary structure would open novel avenues in disease treatment and regulation of cellular processes. In pursuit of this objective, probabilistic models, especially those employing machine learning have come to dominate RNA secondary structure prediction, proving better than previous tools which were based upon comparative sequence analysis or folding algorithms employing thermodynamic and stochastic parameter schemes. This review studies landmark models using machine learning techniques, which have accelerated the research in RNA secondary structure prediction in the past two decades. Additionally, an elaboration on RNA types, functions, and their functional mechanisms in diseases, is intended to provide the reader with the prerequisite knowledge to understand the vitality of unearthing structural information of RNA. An outline of alternate RNA structure tools and techniques is also incorporated for a better understanding of the theme.**

Keywords—RNA secondary structure prediction, pair probabilistic models, machine learning

I. INTRODUCTION

Ribonucleic acid or RNA is an important part of the genetic machinery and serves as a single-stranded repository of genomic data like DNA (double-stranded). It is in the form of a chain formed by nucleotide bases-bound-ribose sugars which are joined by phosphodiester bonds. More importantly, it functions by regulating genes, synthesizing proteins, and micro-managing cellular processes (example: translation). The nucleotide bases in RNA are – adenine, guanine, cytosine, and uracil, which replaces thymine from DNA.

Traditionally RNA functions were thought to be confined to those carried forward by three major RNAs – the mRNA: a product of transcription from the DNA, the rRNA: molecules that build the cell's translational machinery along with proteins i.e. ribosomes and the tRNA: molecules that together with rRNA facilitate the translation of codons in mRNA to amino acids in proteins. In the process of protein synthesis, the messenger RNA (mRNA) is the carrier of genetic code (sequence), obtained by the transcription of the DNA, which are carried to the ribosomes. mRNA is synthesized in the nucleus and it travels to the cytoplasm, where translation takes place. The ribosome, the fundamental unit for translation, is composed of proteins and ribosomal RNA (rRNA). These proteins in ribosome subunits are encoded by the rRNA itself, in the nucleolus of the nucleus. The ribosomal assembly once complete with rRNA and proteins travel to the cytoplasm where they translate the information in mRNA in the form of 'codons' to amino acids. This translation process is essentially facilitated by the soluble or activator RNA which is commonly known as the transfer RNA (tRNA) which are made up of fewer than 100 nucleotides. Codons are short sequences of three nucleotides to which a specific amino acid is linked. As the ribosome traverses through the mRNA, the tRNA merges the amino acids specific to codons into a sequence, to form proteins [1].

However, RNAs function in a much higher capacity, with many more roles – a fact that has been especially entrenched in the last 3 to 4 decades. Apart from regulating the transcription process, from DNA to RNA, the regulation of DNA replication is also carried out by ribonucleic acid chains. RNA molecules also exist as enzymes known as Ribozymes, which carry out biochemical reactions by serving as catalysts. The cause for RNA playing important tasks during the diseased state as well as in normal cellular mechanisms can be attributed to their surfeit in the human body, the diversity of functions they perform and their involvement in a vast number of complex regulatory processes.

Supplementing the rRNA, tRNA and mRNA, the RNA exists in several other forms. Two major classes of mRNA can be identified, namely – coding RNA (cRNA) i.e. the mRNA and noncoding RNA (ncRNA). Noncoding RNA are of two types – the 'housekeeping ncRNA', which include the tRNA and the rRNA; the second type are the 'regulatory ncRNA'. These regulatory ncRNA, depending on their length size can be 'long ncRNA' with a sequence length of more than 200 nucleotides or 'small ncRNA' with a sequence length of fewer than 200 nucleotides. The small regulatory ncRNA have higher significance in terms of functions, it is observed. Small ncRNA are of five types – microRNA (miRNA), small-interfering RNA (siRNA), PIWI-interacting RNA (piRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA). Among these five, the usually 22 nucleotide long miRNA are of great import; they are endogenous and are found in most eukaryotes and function as gene regulators. miRNAs cause gene silencing, they do this by binding to target mRNA and this prevents them from producing functional proteins since translation is inhibited. An increasing number of research studies have established miRNAs as potential biomarkers for the diagnosis of human cancer while their dysregulation can lead to diseases such as Alzheimer's [2], [3]. Further studies have indicated that both miRNAs and siRNAs play a role in gene silencing or RNA interference. While miRNA are endogenous, siRNA are

978-1-7281-6324-6/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

primarily exogenous. The 26-32 nucleotide long piRNA also regulate the expression of the 'jumping genes' (transposons) by preventing these genes from getting transcribed into germ cells. Majority of piRNAs are complementary to specific transposons and this helps in target specificity for transposon silencing. Circular RNAs are circular in shape and are formed when its 3' and 5' end fuse into a loop. They enable protein synthesis but can also bind miRNA, thus preventing miRNAs from carrying out silencing. Besides this circRNA also regulate the process of alternative splicing (from which they are synthesized) and also transcription [4], [5].

RNA plays an especially important role in when it comes to studying, treating or in some cases even facilitating human diseases. An example, as mentioned earlier, is the regulation of genes associated with cancer by a few miRNA, in a manner that expedites tumour development. Furthermore, dysregulation of miRNA has been associated with Alzheimer disease and various other neurodegenerative disorders. tRNA-derived fragments (tRFs) are noncoding RNAs which are being researched as cancer-causing agents [3]. Also, caspases are proteins that cause programmed cell death i.e. apoptosis; tRNA have been found to bind to caspases causing their inhibition, which enables the unhindered proliferation of cells leading to cancer. MALAT1 or 'metastasis-associated lung adenocarcinoma transcript 1' are new groups of RNA transcripts specific to tumours, which have been identified owning to latest RNA sequencing methodologies; their increased levels indicate growth and spread of tumour cells since they are found in high amounts in cancerous tissues [6]. Additionally, some classes of RNA have been found to isolate RNA-binding proteins which accumulate in neural tissues causing neurological diseases such as myotonic dystrophy and ALS (amyotrophic lateral sclerosis) [7].

Unlike DNA which has two complementary strands that add to its stability, RNA has a single strand and is therefore unstable, possesses high energy and is hence, chemically very reactive. Also, adding to its instability is the 2′-hydroxyl group on the ribose ring, whose presence facilitates RNA hydrolysis – a reaction in which the phosphodiester bonds linking ribose and phosphate groups are cleaved, causing the RNA to break. However, RNA is as rapidly being synthesized inside the cell, as it is being degraded. Furthermore, the formation of the three-dimensional structure due to self-folding of RNA increases its stability. These 3D structures are formed because of the internal base-pairing within an RNA molecule or even non-canonical base pair interactions. The self-complementary sequences in RNA enable it to form a variety of complex self-folding structures through canonical and nested base pairing, making hairpin loops, stem-loops, bulges and even helices, internal and external loops, multiloops, pseudoknots, along with non-canonical and non-nested base pairs, including triplet interactions (refer Fig. 1). The three-dimensional structure is essential for the stability of RNA and more importantly, it enables RNA to carry out its regulatory functions more suitably. It also allows for the attachment of chemical groups such as, the methyl group, which further stabilize the RNA secondary structure. For example, a tRNA molecule with a methyl group at 58 the position of the tRNA sequence makes the molecule more stable and functional whereas a tRNA

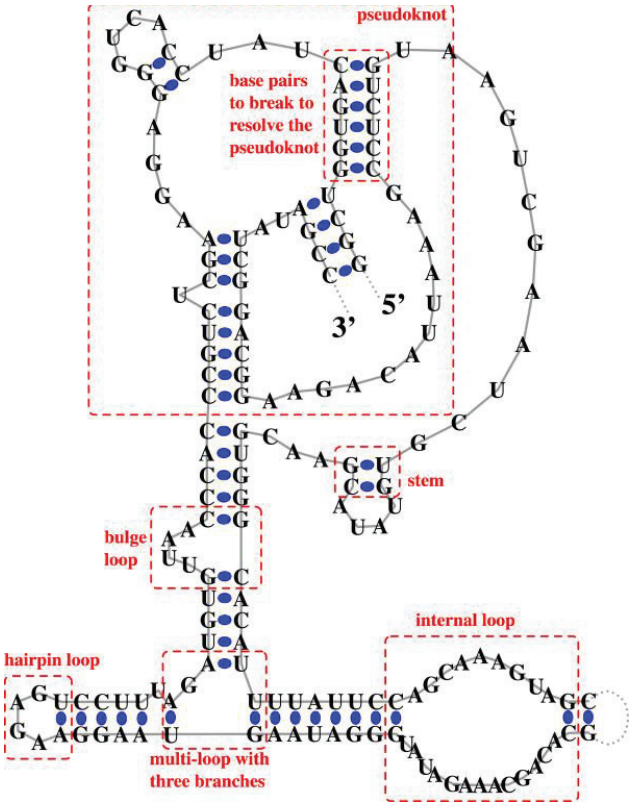


Fig. 1. Diagrammatic representation of self-folded RNA Secondary Structure with numerous types of loops (Image Source: Andronescu M. et al., 2008)

without methyl group becomes non-functional and eventually gets degraded. RNA also exist in conjugation with proteins forming ribonucleoproteins (RNPs); research has shown the RNA part of RNPs function as catalysts, like proteins [8], [9].

After the discovery of complex structures of RNA, it became evident that the free sites in the 3D conformation of RNA molecules are the ones participating in binding to proteins or to other RNA, which then enable the RNA molecules to carry out their varied functions. While RNA were found to exist in the form of tertiary and secondary structures (refer Fig. 2), an understanding dawned that predicting tertiary structures from RNA primary structures is a very complicated task due to the everchanging chaotic cellular environment and the dynamic nature of RNAs. Therefore, the study of RNA secondary structures from their primary structures and sequences came to be embarked upon. However, most of the human RNA molecules are noncoding RNAs whose structures and functions are unknown. Therefore, the need arose and it still persists, to acutely study the secondary structures of RNA, identify the sequence of their bound sites and those of the free sites in order to better study and identify functional mechanisms of RNAs. This requires highly precise base-pairing or secondary structure prediction of RNA.

Accurate base-pairing study of RNA nucleotides in 3D secondary structures, in high resolution of single base pairs, can be performed by the experimental techniques of X-ray crystallography, nuclear magnetic resonance (NMR), or

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

cryogenic electron microscopy. However, these require adept handling by experts along with large experimental time. Added to this the low availability of experimentally determined structures – about <0.01% of the 14 million noncoding RNAs collected in RNAcentral [10] – developing cost-effective and precise computational methods for direct prediction of RNA secondary structure from the sequence is a much more desirable and suitable method.

Strong bonds such as Hydrogen bonds and base stacking result in the formation of RNA secondary structures. The first methods for structure prediction were physical techniques (PARS technique [11], DMS method [12], SHAPE model [13]). The existing prediction methods for these RNA secondary structures can be differentiated into two categories – comparative sequence analysis methods and methods involving folding algorithms with scoring schemes of statistical, thermodynamic, or probabilistic nature [10]. Conserved base pairs among homologous sequences are determined by comparative sequence analysis; availability of a large number of such homologous sequences make this method highly successful and precise (Sankoff method [14], Knudsen & Hein method [15] and Alali & Sagot method [16]. However, only about 3000 families are known in Rfam [17], which is insufficient. Therefore, folding algorithms were developed to predict RNA secondary structure, where an RNA sequence is folded as per suitable scoring function. In one prediction method, RNA structures are divided into loops and stems i.e. substructures, and according to the nearest-neighbour model, the structure with minimum free energy (MFE) or global minimum, combined with the technique of dynamic programming is estimated [10]. Many variations of this method exist, resulting in several tools with thermodynamically determined parameters, such as in RNAshapes [18], RNAstructure [19], RNAfold [20] models. In another prediction method, probabilistic models are employed to obtain parameters upon which the model is based; these scoring functions can either be attained through random probability-based statistical schemes, using stochastic context-free grammars (SCFGs), or they can be obtained through machine learning techniques (deep learning). Some of the famous machine learning-based tools are ContextFold [9], CONTRAfold [21] CentroidFold [22] models.

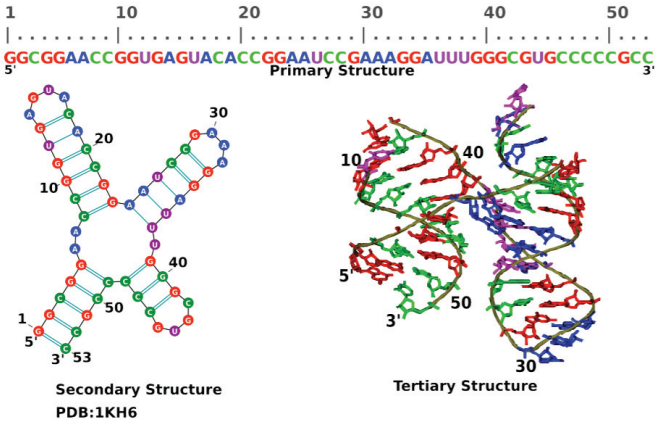


Fig. 2. Representation of primary, secondary and tertiary structures of RNA; PDB:1KH6 (Image Source: Kim N. et al., 2013)

Machine Learning is the branch of Artificial Intelligence in which system (machines) are trained to extract a complex mathematical model from a set of known examples (data), in order to identify patterns and make predictions on a set of unknown examples, with minimal human intervention. Artificial Neural Networks emulate the Biological Neural Networks in the human brain. The length of an ANN can be attributed to the number of hidden layers and the width to the maximum number of neurons in a layer. ANN started being addressed as 'deep neural networks' when it became possible to train ANN models with a large number of hidden layers. CNN or Convolutional Neural Networks are a class of deep neural networks greatly used for image analysis. They employ 'convolution' which is a linear mathematical operation used by a CNN in at least two layers, instead of general matrix multiplication. Recurrent Neural Networks or RNN are types of ANN having memory of previous time step; they are multilayer with each layer representing observations at a certain time period and are called recurrent because they perform same work of every item of a sequence with the current output being dependent on the output of previous time step, which is their main advantage [23].

II. Machine Learning Methods for RNA Secondary Structure Prediction

In the context of RNA secondary structure prediction, machine learning methods are essentially used to obtain scoring parameters of RNA substructures. These methods are gaining increasing popularity due to their ability to suggest a great number of possible structures computationally, with high accuracy in contrast to arduous and thermodynamically derived scoring parameters. Popular RNA secondary structure prediction tools developed using machine learning algorithms are described below.

A. CONTRAfold

Previously RNA secondary structure prediction through parameter based estimation was limited to minimum free energy (Mfold [24]) methods or the later developed stochastic context-free grammar (SCFGs) methods, due to the high accuracy obtained by these two prediction approaches. While the former depended on experimental thermodynamic parameters and gave the most optimal structure for short sequences, the latter utilized statistical and joint probability-based algorithms to derive parameters, though with moderate accuracy which was lesser than the empirical methods. However, RNA in cells, as it has come to be understood, do not necessarily exist in conformations with minimum free energy. Additionally, methods predicting structures of longer RNA sequences and with greater accuracy were required.

CONTRAfold, developed in 2006 became the first and the foremost method to predict RNA secondary structures using conditional log-linear models (CLLMs) with an accuracy greater than thermodynamic methods. CLLMs emerged as the malleable group of probabilistic models which in addition to learning parameters and optimization (like SCFGs), also utilized discriminative training and feature-abundant scoring to include schemes just like those used by energy-dependent thermodynamic models. The model proposed a novel

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

Maximum Expected Accuracy (MEA) estimator to achieve this. More importantly, CONTRAfold achieved the highest prediction accuracy for single sequences at the time, outperforming all physics-dependent and previous probabilistic models, although not accounting for pseudoknots [21].

B. CentroidFold

The earlier Minimum Free Energy (MFE) based models used Maximum likelihood (ML) estimators for prediction, which gave low probabilities of prediction and lack of optimal predictions [24]. Subsequently, the SCFG models considered an ensemble of plausible structures for better prediction accuracy and proposed several new estimators to achieve the same, which was a more holistic approach. These included the 'centroid' estimator, by Sfold [25] and MEA estimator by CONTRAfold [21].

In this regard the approach that the authors of the current model followed was attempted at upgrading previous estimators, to increase accuracy, which was proven both theoretically and experimentally. CentroidFold proposed a new estimator called γ-centroid estimator for a more accurate prediction [22]. As an added advantage, the tool allows the user to input RNA in two forms – either as a single sequence (FASTA format) or as a multiple sequence alignment (CLUSTAL-W format) with other RNA. The method emphasizes the critical importance of posterior techniques for identifying RNA secondary structures. However, CentroidFold also does not factor in pseudoknots in its predictions.

C. ContextFold

With an increasing focus on machine learning-based algorithms for parameter estimation for RNA secondary structure prediction, a constant endeavour to improve these methods endured. It was observed that the number of model parameters in such models remained almost the same over the years and across several tools, even as the algorithms for prediction improved. For this the authors proposed 'rich parameterization' model in ContextFold, where the information utilized for structure prediction was increased, implying an increase in the number of parameters (about 70,000 free parameters).

The added advantage of advanced RNA folding algorithms, increasing availability of RNA sequential [26], [27] and experimentally determined structural data (larger training sets) [28] and progressions in machine learning techniques [29], [30] all contributed to the increasing parameterization. The structural elements defined by the ContextFold tool are founded on those in the Turner99 model [31]. This, in addition to analyzing long sequences, and more structural elements resulted in a novel folding prediction model with a notably higher quality of predictions [9].

D. SPOT-RNA

Machine learning-based probabilistic models before SPOT-RNA were successful in predicting, with good accuracy, the canonical and nested base pairs towards RNA secondary structure prediction. However, the accuracy of prediction could not cross 80% ceiling in any of these. This was because these models ignored base pairing which was stabilized by tertiary interactions such as – unstacked (solitary), non-canonical (that do not follow Watson-Crick base-pairing of A-U, G-C and G-U), and non-nested base pairs in addition to triplet interactions [4], [32].

However, SPOT-RNA is a more holistic model inclusive of all types of base pairing including those specific to tertiary structures. The software employs deep learning techniques with convolutional neural network (CNN) layers, two-dimensional Bidirectional Long Short-Term Memory (2D-BLSTM) layers and merging these with the use of 'transfer learning' from a model trained with bpRNA dataset of higher quality, containing more than 10,000 nonredundant RNA obtained through comparative analysis; this improved upon the limitation of having <250 RNA structures of high resolution – a limitation that deterred previous studies [10]. The model improved not just the accuracy of RNA structure prediction, but also of sequence alignments and functional annotations. The model also improves upon the computation time, even though the sequences used as inputs in training are of limited length.

III. Additional Methods

A. RNA tertiary structure prediction

Less attention was allotted to nucleotide interactions that contributed to the tertiary structure of RNA, in the standard models for secondary structure prediction (except SPOT-RNA). This resulted in a performance ceiling of 80% in those models, which in turn motivated the research community to develop models for secondary structure prediction based on tertiary interactions as additional parameters [10].

TABLE I. Comparison of popular RNA Secondary Structure Prediction Algorithms

Name	Algorithm type	Machine learning used	Pseudoknot Prediction	Ref[a]
Sankoff (1985)	Comparative Sequence Analysis	No	No	[14]
Knudsen & Hein (1999)	Comparative Sequence Analysis	No	No	[15]
Alali & Sagot (2005)	Comparative Sequence Analysis	No	No	[16]
RNAshapes (2014)	Thermodynamic	No	No	[40], [18]
RNAstructure (2010)	Thermodynamic	No	Yes	[19]
RNAfold (2011)	Thermodynamic	No	No	[20]
CONTRAfold (2006)	Probabilistic	Yes	No	[21]
CentroidFold (2009)	Probabilistic	Yes	No	[22]
ContextFold (2011)	Probabilistic	Yes	No	[9]
SPOT-RNA (2019)	Probabilistic	Yes	Yes	[10]
Knotty (2018)	CCJ 2.0	No	Yes	[36]
IPknot (2011)	MEA and Heuristic Integer Programming	No	Yes	[35]
MC-Fold and MC-Sym Pipeline (2008)	MC-Fold and MC-Sym	No	Yes	[38]
CDP-Fold (2019)	Probabilistic and Dynamic Programming	Yes	Yes	[41]

[a] Ref = Reference

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

These models are divided into those that predict secondary structures with pseudoknots (e.g. pknotsRG [33], Probknot [34], IPknot [35], and Knotty [36]), and those that predict secondary structures with non-canonical base pairs (e.g. CycleFold [37], MC-Fold [38], and MCFold-DP [39]).

However, the accuracy of these methods is lower than popular models and none of them considers both pseudoknots and non-canonical base pairs for prediction, and neither does any model consider base triplets and lone base pairs.

B. *Hybrid methods*

Popular predictive models are demarcated according to algorithms – comparative sequence analysis, and folding algorithms (thermodynamic, stochastic or probabilistic). Few attempts have also been made to predict RNA structures by merging two algorithms. For example, CDP-Fold [41] algorithm employs both convolutional neural networks (CNNs) and dynamic programming (MFE). While suggesting modest accuracy, such algorithms require further optimization and incorporation of more parameters to predict RNA secondary structures satisfactorily.

C. *RNA Inverse Folding*

Efforts have also been made towards designing RNA sequences based upon target secondary structures that are required to perform certain functions. Several advanced algorithms for inverse folding of RNA exist, such as downhill simplex algorithm, multi-objective genetic algorithm, local search algorithm, etc. RNA inverse [42], Frnakestein [43], EteRNABot [44], IncaRNAtion [45], RNAiFold [46], and MODENA [47] are among popular inverse folding tools.

IV. Discussion

While machine learning models have proven to be highly effective in predicting RNA secondary structures, more significantly than minimum free energy (MFE) and stochastic context-free grammar (SCFG) methods, they have limitations that must be addressed. Biasness of the model to training data leads to overfitting – which in turn decreases the veracity of the predictions. Overfitting has been known to exist in many models, including ContextFold [48]. A reason for this is also because of the dominance of RNA databases with a single RNA type. The lack of diversity of well-curated single-sequence RNA secondary structures contributes to the problem and prevents improvement in statistical models for RNA secondary structure prediction. Furthermore, there is a longstanding need for a greater number of experimentally identified RNA structures with high-resolution, which are expected to further increase the efficacy of probabilistic models. In retrospect, a futuristic model with sufficient experimentally determined RNA secondary structures and a robust probabilistic model for predicting unknown structures could be ideal. Finally, the universal problem of 'black box' in machine learning models prevents an understanding of how features are extracted till the output layer, which robs one of the potentials of improving the models from a more biological perspective. Use of RNN layers in probabilistic models, specifically LSTMs and GRUs, may offer greater accuracy since they tackle the issue of long-term dependency, which the RNA secondary prediction methods can greatly benefit from.

V. Conclusion and Scope

Prediction of RNA secondary structures would find major applications in several research areas of RNA biology. It would enable prediction of ribosomal-binding sites on mRNA which code for proteins, opening avenues to regulate the process, in the context of disease treatment. It would help determine protein binding sites in RNA-binding proteins, which significantly regulate post-transcription, translation, and splicing. Also, the prediction of accessible sites on RNA secondary structure would imply the identification of miRNA binding sites in Eukaryotes and siRNA binding sites in Prokaryotes. Knowledge of the location of these sites would empower the regulation of RNA interference. It would also contribute to RNA therapy – which involves designing drugs to regulate specific RNA. All these discoveries could potentiate a deeper understanding of RNA-disease relationship, which would lead to better healthcare systems, more adept at combating diseases. Lastly, an improvement in machine learning probabilistic models along with an adequate increase in experimentally obtained secondary structure data, and an identification of the nucleotides involved in pairing, would undoubtedly usher us into an era of multifold advancements in RNA biology.

References

[1] J. Brosius and C. A. Raabe, "What is an RNA? A top layer for RNA classification," *RNA Biol.*, 2016.

[2] K. N. Holohan, D. K. Lahiri, B. P. Schneider, T. Foroud, and A. J. Saykin, "Functional microRNAs in Alzheimer's disease and cancer: Differential regulation of common mechanisms and pathways," *Frontiers in Genetics*. 2013.

[3] C. Battaglia *et al.*, "Candidate genes and miRNAs linked to the inverse relationship between cancer and alzheimer's disease: Insights from data mining and enrichment analysis," *Front. Genet.*, 2019.

[4] J. Nowakowski and I. Tinoco, "RNA structure and stability," *Semin. Virol.*, 1997.

[5] L. Chen and G. Wong, "Transcriptome Informatics," in *Encyclopedia of Bioinformatics and Computational Biology*, Elsevier, 2019, pp. 324–340.

[6] N. Amodio *et al.*, "MALAT1: A druggable long non-coding RNA for targeted anti-cancer approaches," *Journal of Hematology and Oncology*. 2018.

[7] B. Swinnen, W. Robberecht, and L. Van Den Bosch, "RNA toxicity in non-coding repeat expansion disorders," *EMBO J.*, 2020.

[8] K. Darty, A. Denise, and Y. Ponty, "VARNA: Interactive drawing and editing of the RNA secondary structure," *Bioinformatics*, 2009.

[9] S. Zakov, Y. Goldberg, M. Elhadad, and M. Ziv-Ukelson, "Rich parameterization improves RNA structure prediction," *J. Comput. Biol.*, 2011.

[10] J. Singh, J. Hanson, K. Paliwal, and Y. Zhou, "RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning," *Nat. Commun.*, 2019.

[11] M. Kertesz *et al.*, "Genome-wide measurement of RNA secondary structure in yeast," *Nature*, 2010.

[12] Y. Ding, Y. Tang, C. K. Kwok, Y. Zhang, P. C. Bevilacqua, and S. M. Assmann, "In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features," *Nature*, 2014.

[13] K. A. Wilkinson *et al.*, "High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states," *PLoS Biol.*, 2008.

[14] D. Sankoff, "SIMULTANEOUS SOLUTION OF THE RNA FOLDING ALIGNMENT AND PROTOSEQUENCE PROBLEMS.," *SIAM J. Appl. Math.*, 1985.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

[15] B. Knudsen and J. Hein, "RNA secondary structure prediction using stochastic context-free grammars and evolutionary history," *Bioinformatics*, 1999.

[16] J. Allali and M. F. Sagot, "A new distance for high level RNA secondary structure comparison," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2005.

[17] I. Kalvari *et al.*, "Non-Coding RNA Analysis Using the Rfam Database," *Curr. Protoc. Bioinforma.*, 2018.

[18] S. Janssen and R. Giegerich, "The RNA shapes studio," *Bioinformatics*, 2015.

[19] J. S. Reuter and D. H. Mathews, "RNAstructure: Software for RNA secondary structure prediction and analysis," *BMC Bioinformatics*, 2010.

[20] R. Lorenz *et al.*, "ViennaRNA Package 2.0," *Algorithms Mol. Biol.*, 2011.

[21] C. B. Do, D. A. Woods, and S. Batzoglou, "CONTRAfold: RNA secondary structure prediction without physics-based models," in *Bioinformatics*, 2006.

[22] K. Sato, M. Hamada, K. Asai, and T. Mituyama, "CentroidFold: A web server for RNA secondary structure prediction," *Nucleic Acids Res.*, 2009.

[23] S. Min, B. Lee, and S. Yoon, "Deep learning in bioinformatics," *Briefings in bioinformatics*, vol. 18, no. 5. pp. 851–869, 2017.

[24] M. Zuker, "Mfold web server for nucleic acid folding and hybridization prediction," *Nucleic Acids Res.*, 2003.

[25] Y. Ding, C. Y. Chan, and C. E. Lawrence, "Sfold web server for statistical folding and rational design of nucleic acids," *Nucleic Acids Res.*, 2004.

[26] Y. Wexler, C. Zilberstein, and M. Ziv-Ukelson, "A study of accessible motifs and RNA folding complexity," in *Journal of Computational Biology*, 2007.

[27] R. Backofen, D. Tsur, S. Zakov, and M. Ziv-Ukelson, "Sparse RNA folding: Time and space efficient algorithms," in *Journal of Discrete Algorithms*, 2011.

[28] S. Griffiths-Jones, S. Moxon, M. Marshall, A. Khanna, S. R. Eddy, and A. Bateman, "Rfam: Annotating non-coding RNAs in complete genomes," *Nucleic Acids Res.*, 2005.

[29] M. Collins, "Discriminative training methods for hidden Markov models," 2002.

[30] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz, and Y. Singer, "Online passive-aggressive algorithms," *J. Mach. Learn. Res.*, 2006.

[31] R. Kierzek, M. E. Burkard, and D. H. Turner, "Thermodynamics of single mismatches in RNA duplexes," *Biochemistry*, 1999.

[32] E. Westhof and V. Fritsch, "RNA folding: Beyond Watson-Crick pairs," *Structure*. 2000.

[33] J. Reeder and R. Giegerich, "Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics," *BMC Bioinformatics*, 2004.

[34] S. Bellaousov and D. H. Mathews, "ProbKnot: Fast prediction of RNA secondary structure including pseudoknots," *RNA*, 2010.

[35] K. Sato, Y. Kato, M. Hamada, T. Akutsu, and K. Asai, "IPknot: Fast and accurate prediction of RNA secondary structures with pseudoknots using integer programming," *Bioinformatics*, 2011.

[36] H. Jabbari, I. Wark, C. Montemagno, and S. Will, "Knotty: Efficient and accurate prediction of complex RNA pseudoknot structures," *Bioinformatics*, 2018.

[37] M. F. Sloma and D. H. Mathews, "Base pair probability estimates improve the prediction accuracy of RNA non-canonical base pairs," *PLoS Comput. Biol.*, 2017.

[38] M. Parisien and F. Major, "The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data," *Nature*, 2008.

[39] C. Höner zu Siederdissen, S. H. Bernhart, P. F. Stadler, and I. L. Hofacker, "A folding algorithm for extended RNA secondary structures," *Bioinformatics*, 2011.

[40] Z. J. Lu, J. W. Gloor, and D. H. Mathews, "Improved RNA secondary structure prediction by maximizing expected pair accuracy," *RNA*, 2009.

[41] H. Zhang *et al.*, "A new method of RNA secondary structure prediction based on convolutional neural network and dynamic programming," *Front. Genet.*, 2019.

[42] I. L. Hofacker, "Vienna RNA secondary structure server," *Nucleic Acids Res.*, 2003.

[43] R. B. Lyngsø, J. W. J. Anderson, E. Sizikova, A. Badugu, T. Hyland, and J. Hein, "Frnakenstein: Multiple target inverse RNA folding," *BMC Bioinformatics*, 2012.

[44] J. Lee *et al.*, "RNA design rules from a massive open laboratory," *Proc. Natl. Acad. Sci. U. S. A.*, 2014.

[45] V. Reinharz, Y. Ponty, and J. Waldispühl, "A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotide distribution," in *Bioinformatics*, 2013.

[46] J. A. Garcia-Martin, P. Clote, and I. Dotu, "RNAiFOLD: A constraint programming algorithm for rna inverse folding and molecular design," *J. Bioinform. Comput. Biol.*, 2013.

[47] A. Taneda, "Multi-objective optimization for RNA design with multiple target secondary structures," *BMC Bioinformatics*, 2015.

[48] E. Rivas, R. Lang, and S. R. Eddy, "A range of complex probabilistic models for RNA secondary structure prediction that includes the nearest-neighbor model and more," *RNA*, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on December 04,2020 at 05:52:55 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

Robust Stator Winding Fault Detection in PMSMs with Respect to Current Controller Bandwidth

Shaopo Huang, Anmol Aggarwal, *Student Member, IEEE*, Elias G. Strangas, *Senior Member, IEEE*, Kui Li, Feng Niu, *Member, IEEE*, and Xiaoyan Huang, *Member, IEEE*

***Abstract*—Permanent magnet synchronous machines (PMSMs) work with controllers and controller parameters affect the machine's voltages and currents and thus the diagnosis of stator winding fault. One such parameter is the current controller bandwidth that affects the harmonic content in these signals. In the literature, the 2nd harmonics in both dq-axis currents and voltages have been used for fault detection, and the main advantage is that inter-turn short-circuit (ITSC) winding fault with a few shorted turns, specially one turn, can be detected. One indicator that combines the fault information in voltages and currents is the instantaneous reactive power (IRP). However, it has been shown that IRP does not perform robustly with respect to current controller bandwidth. The main contribution of this work is that it proposes and validates a new mathematical function for ITSCs detection, Rayleigh quotient, which combines the fault information present in voltages and currents to make ITSCs detection robust to current controller bandwidth. The validations from both co-simulation and experiment show that the proposed method is robust to the bandwidth and can detect ITSC faults faster than IRP.**

***Index Terms*—controller bandwidth, fault detection, inter-turn short-circuit, PMSM, Rayleigh quotient, 2nd harmonic.**

I. Introduction

Winding faults due to insulation degradation and failure between turns, leading to short circuits (stator inter-turn short-circuits, ITSCs) are among the most common faults in permanent magnet synchronous machines (PMSMs) that lead to catastrophic failures in short time [1]–[3]. The problems of fault detection during operation has been addressed extensively over the years. This detection is based on the controller characteristics, variables and operation [4]–[7] and has led to significant improvements. A number of issues remain to be addressed: sensitivity to the fault severity, accuracy, speed of detection, so that timely action can be taken, and robustness to the characteristics of the controller.

Motor Current Signature Analysis (MCSA) and Motor Voltage Signature Analysis (MVSA) are effective techniques for faults detection [2], [3], [8]–[11]. These methods are non-invasive and the signals, currents and voltages, are readily available in the controller. In synchronous frame of reference, ITSCs change the dc components of the voltages, and introduce 2nd harmonics. The shifts of the dc components in commanded dq-axis voltages could be used to detect the ITSC with 5% shorted turns of phase coils [3]. A stator winding interturn fault across 50% of phase turns was detected by extracting the magnitudes of the 2nd harmonics in reference voltages [5], but the detecting performance with low percentage of phase winding shorted was not studied in detail. ITSCs in PMSMs were detected by both the dc components and the 2nd harmonics in reference voltages in [12]. On the other hand, ITSCs also introduce characteristic components into stator currents. The interturn faults could be diagnosed by analyzing the 3rd harmonics in stator currents, and the detectable percentage of shorted turns was as low as 2-3%, regardless of machine winding configuration [13]. In [14], the severity of ITSC was determined using the amplitude of the side band components of stator currents, and the frequency component in the magnitude of extended Park vector were employed to detect the fault. Amplitudes of the first 15 harmonics (i.e., the fundamental and the 2nd-15th harmonics) of measured currents or commanded voltages were used as detailed features for faults detection and classification in [15]. Through linear discriminant analysis, the ITSCs across 12% of phase turns for a distributed windings permanent magnet (PM) machine and that of 4.4% for a concentrated windings PM machine could be detected and separated from other faults.

The techniques discussed above are mainly based on analysis of one signal, voltages or currents. However, the low severity of incipient faults with few shorted turns makes it difficult to recognize them. For example, in the same machine analyzed in this work, the one-turn short-circuit is undetectable experimentally through the method in [3]. Moreover, an over-current or differential relay in power systems cannot protect the machine with ITSC, because the current sensors are placed outside the stator and unable to measure the circulating current. The effects of stator interturn faults on the line

Manuscript received xxx xx, 20xx; revised xxx xx, 20xx; accepted xxx xx, 20xx. Date of publication xxx xx, 20xx; date of current version xxx xx, 20xx. This work was supported in part by the State Scholarship Fund of China under Grant 201606700009, in part by the National Natural Science Foundation of China under Grant 51707174 and 52077054, and in part by the Natural Science Foundation of Hebei Province under Grant E2019202092. (*Corresponding author: Xiaoyan Huang*)

S. Huang, K. Li and F. Niu are with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300401, China (e-mail: shaopo33@gmail.com; likui@hebut.edu.com; niu.feng.hebut@gmail.com).

A. Aggarwal and E. G. Strangas are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: aggarw43@egr.msu.edu; strangas@egr.msu.edu).

X. Huang is with the College of Electrical Engineering, Zhejiang University, Hangzhou 310058, China (e-mail: xiaoyanhuang@zju.edu.cn).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier xxx/TPEL.xxx

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

currents are often limited [13], [14]. Therefore, indicators were developed based on the fault information in both currents and voltages to detect stator winding faults, even for low severity.

The harmonic fault information and the degree of division of this information between current and voltage varies with the bandwidth of speed and current controllers, and these effects have to be investigated if MCSA (also MVSA) is used for fault detection [16]. This characteristic was also mentioned in [5], [6], [12], [13] and [15]. However, the detailed analytical and physical insights into it were not reported, and the effects of current controller bandwidth on fault detection were not considered. Therefore, the ITSCs detection methods based on MCSA or MVSA, introduced in [5], [6], [12]–[14] and [15], might not be robust enough with respect to the bandwidth.

One combination of the harmonics in currents and voltages for ITSCs detection is instantaneous real and reactive power [1], [8], [17]–[19]. In synchronous frame of reference, instantaneous reactive power, IRP $= 1.5(v_q i_d - v_d i_q)$, has been reported as an effective indicator for ITSCs detection, since it contains more fault information than current and voltage individually. It was shown that the 2nd harmonic in currents decreases with the increase of current controller bandwidth, while the opposite trends is in voltages. Consequently, the 2nd harmonic in IRP is almost not affected, and two bandwidths were selected to confirm this [8]. Regardless of the bandwidth, IRP is a fixed combination of currents and voltages. Because the harmonics in currents and voltages do not change linearly with the bandwidth, the ITSCs detection based on IRP does not perform robustly with respect to current controller bandwidth, as IRP does not adapt to the bandwidth. Therefore, a function combining the information in voltages and currents, and robust to current controller bandwidth is desired.

The main contribution of this work is that the fault information in currents and voltages is weighted adaptively to the change of current controller bandwidth, and this way an ITSC indicator, incorporating information in both currents and voltages and robust to the bandwidth, is obtained. In this work, the machine is considered to operate in current controlled, rotor field oriented mode, supplied from a voltage source inverter. The variations of the 2nd harmonic in dq-axis currents and voltages along with the bandwidth are analyzed. Then, with respect to the bandwidth, a robust indicator based on Rayleigh quotient function is proposed. The robustness of proposed indicator is validated by co-simulations and experiments with different fault severities at different operating conditions. ITSCs with only one shorted turn can be detected fast and reliably under a wide bandwidth range. The detection speed of proposed indicator is also compared to that of the indicator based on IRP.

The proposed indicator was first introduced in [7], where only one-turn incipient ITSC in co-simulation was presented. In this work, experiment validation is conducted and the effectiveness of proposed indicator is investigated further with one-turn clean ITSC and ITSC with more shorted turns. Moreover, the robustness with respect to current controller bandwidth is evaluated and validated by coefficient of variation.

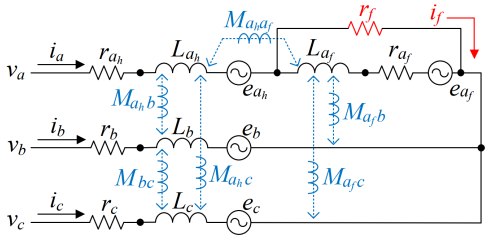


Fig. 1. Three phase electrical model for a PMSM with ITSC.

II. PMSM Model with ITSC

The equivalent circuit model, shown in Fig. 1, has been widely used to study the effect of ITSC. Each phase is modeled using its resistance (r_s), self and mutual inductances (L and M) and the induced back electromotive force (EMF, e_{abc}). An ITSC is introduced in phase A by a fault resistor, r_f, which represents the insulation degradation between turns, and a circulating current, i_f, is generated. The fault begins with relatively high r_f and quickly becomes more severe as the winding insulation deteriorates, potentially forming a clean short-circuit or r_f = 0. Phase A is thus split into two separate windings, denoted as a_h for the healthy part and a_f for the shorted portion.

The machine used to test the proposed method is a single-layer fractional-slot concentrated-winding (FSCW) PMSM with $\frac{2}{5}$ slots per pole per phase, which provides magnetic, thermal and physical separation of phases, and thus the mutual inductances between phases, $M_{a_h b}$, $M_{a_h c}$, $M_{a_f b}$, $M_{a_f c}$ and M_{bc}, are negligible [3], [10], [11]. As presented in [3] and [11], by using Extended Park Transform [20]–[22], the PMSM model with ITSC in the dq reference frame is given as:

$$\begin{aligned} v_d &= r_s i_d + L_d \frac{di_d}{dt} + e_d \\ &\quad - \frac{2}{3}\left(r_{a_f} i_f \sin\theta_e - L_f \sin\theta_e \frac{di_f}{dt} - \omega_e L_f i_f \cos\theta_e \right) \end{aligned} \tag{1}$$

$$\begin{aligned} v_q &= r_s i_q + L_q \frac{di_q}{dt} + e_q \\ &\quad - \frac{2}{3}\left(r_{a_f} i_f \cos\theta_e - L_f \cos\theta_e \frac{di_f}{dt} + \omega_e L_f i_f \sin\theta_e \right) \end{aligned} \tag{2}$$

where the voltage, current, inductance and back EMF of d-axis are denoted as v_d, i_d, L_d and e_d, respectively; and those of q-axis are denoted as v_q, i_q, L_q and e_q. ω_e and θ_e are the electrical angular speed and rotor angle. The resistance of a_f is $r_{af} = \frac{N_f}{N_s} \cdot r_s$, where the number of turns per phase is N_s and that of the fault is N_f. $L_f = -(L_{a_f} + M_{a_h a_f})$, where L_{a_f} is the self inductance of a_f, and $M_{a_h a_f}$ is the mutual inductance between a_h and a_f.

The phase back EMFs in FSCW PMSMs typically include significant 3rd harmonic content. Therefore, in the case of an interturn winding fault, i_f also contains the 3rd harmonic [3]. To evaluate the characteristic of the second part related to the fault in (1) and (2), the circulating current, without considering the higher order harmonics, is modeled as:

$$i_f = I_{f1} \sin(\theta_e + \phi_{f1}) + I_{f3} \sin(3\theta_e + \phi_{f3}) \tag{3}$$

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

where the amplitude and angle of the fundamental of i_f are I_{f1} and ϕ_{f1}, respectively; and that of the 3rd harmonic are I_{f3} and ϕ_{f3} [3], [5]. Substituting (3) into (1) and (2), the d- and q-axis voltages with ITSC become:

$$\begin{aligned} v_d = {} & r_s i_d + L_d \frac{di_d}{dt} + e_d - \frac{1}{3}\{r_{a_f} I_{f1}\cos\phi_{f1} - \\ & r_{a_f}\left[I_{f1}\cos(2\theta_e+\phi_{f1}) - I_{f3}\cos(2\theta_e+\phi_{f3})\right] - \\ & 2\omega_e L_f\left[I_{f1}\sin(2\theta_e+\phi_{f1}) - I_{f3}\sin(2\theta_e+\phi_{f3})\right] - \\ & r_{a_f} I_{f3}\cos(4\theta_e+\phi_{f3}) - 4\omega_e L_f I_{f3}\sin(4\theta_e+\phi_{f3})\} \end{aligned} \tag{4}$$

$$\begin{aligned} v_q = {} & r_s i_q + L_q \frac{di_q}{dt} + e_q - \frac{1}{3}\{r_{a_f} I_{f1}\sin\phi_{f1} + \\ & r_{a_f}\left[I_{f1}\sin(2\theta_e+\phi_{f1}) + I_{f3}\sin(2\theta_e+\phi_{f3})\right] - \\ & 2\omega_e L_f\left[I_{f1}\cos(2\theta_e+\phi_{f1}) + I_{f3}\cos(2\theta_e+\phi_{f3})\right] + \\ & r_{a_f} I_{f3}\sin(4\theta_e+\phi_{f3}) - 4\omega_e L_f I_{f3}\cos(4\theta_e+\phi_{f3})\} \end{aligned} \tag{5}$$

The models (4) and (5) show that ITSC introduces additional dc and oscillating components, the 2nd and 4th order harmonics. The fundamental of i_f determines the additional dc components in the faulty dq-axis voltages. Together with the 3rd harmonic in i_f, it also contributes to the 2nd and the 4th harmonics in the dq-axis voltages.

For the same machine analyzed in this work, simulation and experiment results show that the method based on the change of dc components in v_d and v_q cannot detect ITSC with only one shorted turn, because r_{a_f} is low. To solve this problem, the harmonics in v_{dq} are used for ITSCs detection.

However, as discussed in Section I, the harmonics in voltages and currents depend not only on the machine characteristics, but also on the controller parameters. In most cases, harmonics are present both in currents and voltages. For current controllers with limited bandwidth, in ideal symmetric three-phase machines without neutral conductor, the 3rd harmonic currents are absent because they are in phase with each other and there is no return path. ITSC causes an unbalanced condition [23], [24], in which the 3rd harmonic voltages in three phases change in both amplitude and phase angle, which means that the 3rd harmonic voltages are no longer in phase [7]. If current controllers were perfect, there would be no harmonics in phase currents. However, with limited current controller bandwidth, the imbalance in 3rd harmonic voltages generates the 3rd harmonics in three phase currents. Current PI controllers with high bandwidth suppress the harmonics in currents but increase the harmonics in voltages. In this work, the 2nd harmonics in v_{dq} and i_{dq} are used for fault detection, and the effects of current controller bandwidth on these harmonics are analyzed in the next section.

III. Effects of Current Controller Bandwidth on Currents and Voltages

The ac stator variables are transformed to dc through Park transformation. In the synchronous frame, the design methodology of a PI controller for dc machine can be applied to the design of an ac current controller. If the inertia of the machine and load is large enough, the machine can be modeled as an R-L load with constant disturbance voltage [25], e_d and e_q, as shown in Fig. 2, where i_d^* and i_q^* are the commanded

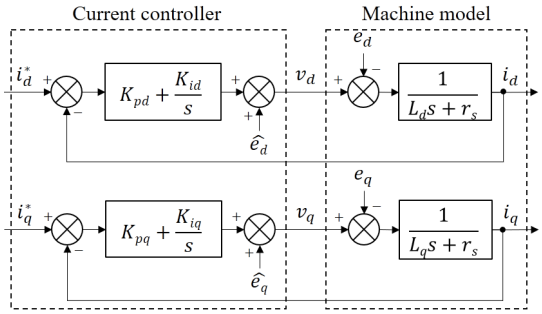


Fig. 2. Current controller diagram in synchronous frame of reference.

values of i_d and i_q. By feed-forward compensation, e_d and e_q are estimated as (6) and the disturbance of EMF can be rejected perfectly. In addition to this rejection, by canceling the pole of the system coming from R-L, with a zero of PI controller, the transfer function of the closed-loop system can be set as first-order low-pass filter. Under this gain tuning method, the proportional and integral gains can be set as (7), and then the closed-loop transfer function of the drive system is reduced to (8), where ω_c is the bandwidth of current PI controller.

$$\widehat{e_d} = -\omega_e L_q i_q, \quad \widehat{e_q} = \omega_e(L_d i_d + \lambda_{pm}) \tag{6}$$

$$K_{pd} = L_d w_c, \quad K_{pq} = L_q w_c, \quad K_{id} = K_{iq} = r_s w_c \tag{7}$$

$$\frac{i_d(s)}{i_d^*(s)} = \frac{i_q(s)}{i_q^*(s)} = \frac{\omega_c}{s+\omega_c} \tag{8}$$

From Fig. 2, the d- and q-axis voltages controlled by the closed-loop current PI controllers can be written as:

$$v_d = (i_d^* - i_d)(K_{pd} + \frac{K_{id}}{s}) + \widehat{e_d} \tag{9}$$

$$v_q = (i_q^* - i_q)(K_{pq} + \frac{K_{iq}}{s}) + \widehat{e_q} \tag{10}$$

Taking the voltage equations (4) and (9), (5) and (10), two differential equations are derived, in which the 4th harmonics are neglected.

$$\begin{aligned} & L_d\frac{d^2 i_d}{dt^2} + (r_s + K_{pd})\frac{di_d}{dt} + K_{id} i_d = K_{id} i_d^* + \\ & \frac{2\omega_e}{3}\{r_{a_f}\left[I_{f1}\sin(2\theta_e+\phi_{f1}) - I_{f3}\sin(2\theta_e+\phi_{f3})\right] \\ & - 2\omega_e L_f\left[I_{f1}\cos(2\theta_e+\phi_{f1}) - I_{f3}\cos(2\theta_e+\phi_{f3})\right]\} \end{aligned} \tag{11}$$

$$\begin{aligned} & L_q\frac{d^2 i_q}{dt^2} + (r_s + K_{pq})\frac{di_q}{dt} + K_{iq} i_q = K_{iq} i_q^* + \\ & \frac{2\omega_e}{3}\{r_{a_f}\left[I_{f1}\cos(2\theta_e+\phi_{f1}) + I_{f3}\cos(2\theta_e+\phi_{f3})\right] \\ & + 2\omega_e L_f\left[I_{f1}\sin(2\theta_e+\phi_{f1}) + I_{f3}\sin(2\theta_e+\phi_{f3})\right]\} \end{aligned} \tag{12}$$

At constant operating conditions, the parameters of PMSM are constant and can be represented by average values. So (11) and (12) can be treated as second order linear differential equations for i_d and i_q, respectively. Then the 2nd harmonics in currents, denoted as i_{d2} and i_{q2}, are derived as:

$$i_{d2} = \frac{2\omega_e Z_f F_d}{3X_d}\sin(2\theta_e + \phi_o + \phi_{d1}) \tag{13}$$

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

$$i_{q2} = \frac{2\omega_e Z_f F_q}{3X_q} \sin(2\theta_e + \phi_o' + \phi_{q1}) \quad (14)$$

where

$$Z_f = \sqrt{r_{a_f}^2 + 4\omega_e^2 L_f^2}$$
$$F_d = \sqrt{I_{f1}^2 + I_{f3}^2 - 2I_{f1}I_{f3}\cos(\phi_{f1} - \phi_{f3})}$$
$$F_q = \sqrt{I_{f1}^2 + I_{f3}^2 + 2I_{f1}I_{f3}\cos(\phi_{f1} - \phi_{f3})}$$
$$X_d = \sqrt{(4L_d^2\omega_e^2 + r_s^2)(4\omega_e^2 + w_c^2)}$$
$$X_q = \sqrt{(4L_q^2\omega_e^2 + r_s^2)(4\omega_e^2 + w_c^2)}$$

In (13) and (14), $\phi_o = \arctan(\frac{-2\omega_e L_f}{r_{a_f}})$, is the impedance angle of the short-circuit path, and $\phi_o' = \phi_o + \frac{\pi}{2}$. The phase angles ϕ_{d1} and ϕ_{q1} are related to the machine parameters and the current controller bandwidth. Substituting (13) and (14) into (9) and (10), respectively, the expressions of the 2nd harmonic voltages can be derived as:

$$v_{d2} = |i_{d2}|(\frac{K_{id}}{2\omega_e}\cos\delta_d - K_{pd}\sin\delta_d) - \omega_e L_q |i_{q2}|\sin\delta_q \quad (15)$$

$$v_{q2} = |i_{q2}|(\frac{K_{iq}}{2\omega_e}\cos\delta_q - K_{pq}\sin\delta_q) + \omega_e L_d |i_{d2}|\sin\delta_d \quad (16)$$

where $|i_{d2}|$ and $|i_{q2}|$ are the amplitudes of i_{d2} and i_{q2}, respectively; $\delta_d = 2\theta_e + \phi_o + \phi_{d1}$ and $\delta_q = 2\theta_e + \phi_o' + \phi_{q1}$.

From (13) and (14), it can be seen that ITSC induces the 2nd harmonics in dq-axis currents, and their amplitudes are affected by the impedance of short-circuit path, circulating current, speed, machine parameters and current controller bandwidth. Current PI controller regulates the current to the commanded value, which appears as constant, but with a limited bandwidth, the current cannot be adjusted to the commanded value perfectly. Instead, harmonic components will exist. As bandwidth increases, which means faster controller with higher gains, the controller adjusts the current to the commanded value more closely by reducing the harmonics, which leads to the increase in the harmonic components in voltage. This is also true for the additional harmonics produced by the winding faults. Therefore, $|i_{d2}|$ and $|i_{q2}|$ decrease as ω_c increases. The trends are shown in Fig. 3, which is drawn according to (7) and (13)-(16) and applied the typical machine used in this work and discussed later. It shows that the opposite trends are seen in the amplitudes of v_{d2} and v_{q2} (marked as $|v_{d2}|$ and $|v_{q2}|$, respectively). The values of r_{a_f}, L_{a_f} and $M_{a_h a_f}$ are determined in accordance with [3], [10] and the machine specifications. The magnitude and angle of the current vector are denoted as I_s and δ, respectively, and ω_m is the machine mechanical speed.

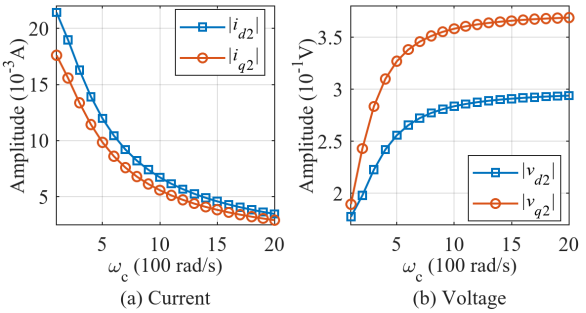


Fig. 3. The amplitudes of the 2nd harmonics under different ω_c, from 100 to 2000 rad/s with a step of 100, at I_s = 10 A with δ = 110° and ω_m = 300 rpm. Fault severity: 1 turn shorted by r_f = 35 mΩ.

The behavior of the 2nd harmonics in currents and voltages shows the need to investigate the effect of current controller bandwidth when designing techniques for ITSCs detection.

IV. Proposed ITSC Indicator

This section proposes a new ITSC indicator, based on Rayleigh quotient function and combining the 2nd harmonics in currents and voltages, and demonstrates that it is robust to current controller bandwidth. First, Rayleigh quotient function is introduced as a weighting rule.

A. Rayleigh Quotient

For a given complex Hermitian matrix $\boldsymbol{X}$ and nonzero vector $\boldsymbol{w}$, the Rayleigh quotient function $R(\boldsymbol{X}, \boldsymbol{w})$ is defined as:

$$R(\boldsymbol{X}, \boldsymbol{w}) = \frac{\boldsymbol{w}^* \boldsymbol{X} \boldsymbol{w}}{\boldsymbol{w}^* \boldsymbol{w}} \quad (17)$$

For real matrices and vectors, the condition of being Hermitian reduces to that of being symmetric, and the conjugate transpose $\boldsymbol{w}^*$ to the transpose $\boldsymbol{w}'$. A symmetric matrix $\boldsymbol{X}$ can be represented as a data matrix $\boldsymbol{A}$ pre-multiplied by its transpose $\boldsymbol{A}'$. Assuming four variables are used for weighting, the data matrix and weight vector are given as $\boldsymbol{A} = [a_1\ a_2\ a_3\ a_4]$ and $\boldsymbol{w}' = [w_1\ w_2\ w_3\ w_4]$, respectively. To make it simple and meaningful as a weighting rule, Rayleigh quotient function is subjected to the constraint of $\boldsymbol{w}'\boldsymbol{w} = \|\boldsymbol{w}\|^2 = 1$, and it becomes:

$$R(\boldsymbol{A}'\boldsymbol{A}, \boldsymbol{w}) = \|w_1 a_1 + w_2 a_2 + w_3 a_3 + w_4 a_4\|^2 \quad (18)$$

Then it is clear that the Rayleigh quotient becomes the square of a weighted average of these four variables.

B. ITSC Indicator Based on Rayleigh Quotient

The 2nd harmonics in dq-axis currents and voltages are not weighted together directly. Instead, the ratios of the 2nd harmonics in i_{dq} and v_{dq} after ITSC to those before the fault, here called current and voltage ratios, are weighted. The ratios of i_d, i_q, v_d and v_q are denoted as a_{i_d}, a_{i_q}, a_{v_d} and a_{v_q}, respectively. Then the ratio matrix, $\boldsymbol{X}$, in the Rayleigh quotient function is:

$$\boldsymbol{X} = \boldsymbol{A}'\boldsymbol{A} \text{ and } \boldsymbol{A} = [a_{id} \quad a_{iq} \quad a_{vd} \quad a_{vq}] \quad (19)$$

From (18) and (19), it is shown that the Rayleigh quotient becomes the square of the weighted average of the current and voltage ratios. Since the 2nd harmonics in currents and voltages vary with current controller bandwidth, the weights should also be different under different bandwidths.

One interesting characteristic of Rayleigh quotient is that, for a given matrix $\boldsymbol{X}$, the function reaches its maximum value, λ_{max}, which is also the largest eigenvalue of $\boldsymbol{X}$, when $\boldsymbol{w}$ is the

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

corresponding eigenvector, $\boldsymbol{v}_{max}$. Similarly, $R(\boldsymbol{X}, \boldsymbol{w}) \geq \lambda_{min}$ and $R(\boldsymbol{X}, \boldsymbol{v}_{min}) = \lambda_{min}$. In this work, the maximum value of Rayleigh quotient, marked as Ray in (20), is used, and the ratio of Ray in faulty condition to that in healthy condition, here called Ray ratio, is proposed as the ITSCs indicator.

$$\text{Ray} = R(\boldsymbol{X}, \boldsymbol{v}_{max}) = \lambda_{max} \tag{20}$$

The current and voltage ratios in healthy condition are 1 and the eigenvalues of $\boldsymbol{X}$ are 0, 0, 0 and 4. Then the maximum eigenvalue is 4 and the corresponding eigenvector is $\boldsymbol{w}' = [0.5\ 0.5\ 0.5\ 0.5]$. The healthy Ray can be calculated using (18) and this $\boldsymbol{w}$. However, according to the characteristic of Rayleigh quotient, Ray is the same as the largest eigenvalue of $\boldsymbol{X}$, on the condition that $\boldsymbol{w}$ is the corresponding eigenvector. Hence Ray can be obtained directly from the online computation of eigenvalue of $\boldsymbol{X}$, and the weight vector, $\boldsymbol{w}$, is not necessary. This means that as the bandwidth changes, the current and voltage ratios change, but they are weighted in Rayleigh quotient function adaptively to this change to give the maximum value, and from which the proposed indicator can be obtained. Compared to the fixed combination of voltages and currents in IRP, the adaptive weights of voltages and current in the Rayleigh quotient make the combination of voltages and currents robust to current controller bandwidth.

The 2nd harmonics in i_d, i_q, v_d and v_q are collected in healthy case under different operating conditions, and stored in a lookup table. Then these harmonics are extracted and the four ratios, a_{i_d}, a_{i_q}, a_{v_d} and a_{v_q}, are calculated online. As discussed above, the healthy Ray is 4. In faulty case, the current and voltage ratios increase and Ray becomes larger than 4, and the ITSC can be detected. The technique of the 2nd harmonic extracting in this work is the same as that in [8], and moving average digital filters were added to reduce the oscillations in Ray.

V. Co-simulation and Experimental Validation

The proposed ITSC indicator is tested through co-simulations (Maxwell and Simplorer) and experiments on a FSCW PM ac machine with design parameters in Table I. The co-simulation combining finite element analysis (FEA) and state variable models, shown in Fig. 4, is better than using simulations with either one individually. The phase resistance, r_s, and the end-winding leakage inductance, L_o, are shown in each phase. The phase A winding is divided into two parts by shorting a portion of it using r_f. The ITSC is introduced by a switch SW.

To introduce ITSC to PMSM in experiment, the phase A winding is divided into separate coils as shown in Fig. 5. The

TABLE I
Machine Specifications

Motor parameters	value	Motor parameters	value
Number of Phases	3	Number of Slots/Poles	12/10
Rated phase current I_{rms}	18 A	Turns per Phase N_s	300
Rated line voltage V_{rms}	480 V	Phase Resistance r_s	1.5 Ω
Base RPM	750	Phase Inductance L_s	112 mH
Magnet flux λ_{pm}	0.28 Wb	Mutual Inductance M_s	≈ 0 H

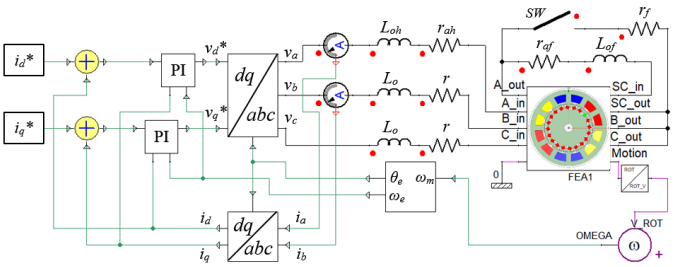


Fig. 4. Co-simulation of PMSM drive with ITSC.

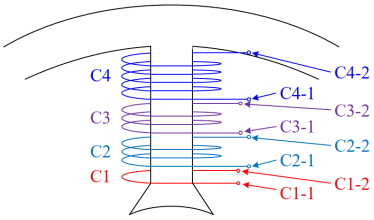


Fig. 5. Stator winding configuration of phase A.

leads of these coils, from C1-1 to C4-2, are brought out of the housing and connected to an interconnection box, where it is possible to externally short the winding and monitor the circulating current. These separate coils are connected in series, and ITSCs with different shorted turns are available by shorting different coils. The experimental setup consists of the tested PMSM, a Dyno operating in constant speed mode, an APS inverter (IAP300T120 PowerStack) and a Real-Time LabVIEW controller, as shown in Fig. 6. The short resistor and the switch for introducing ITSC, SW, are located in a box connected with the PMSM by cables. The controller is developed first in a host computer then deployed to a target computer, and it is run at 40 MHz by a user-programmable FPGA (PCIe-7852R, National Instrument) installed in the target. The PWM switching frequency is 10 kHz, and the controller samples the current and updates the commanded voltage to the PWM switching once every switching period. The three phase currents are measured by current transducers (HAS 50-S, LEM) based on Hall effect measuring principle,

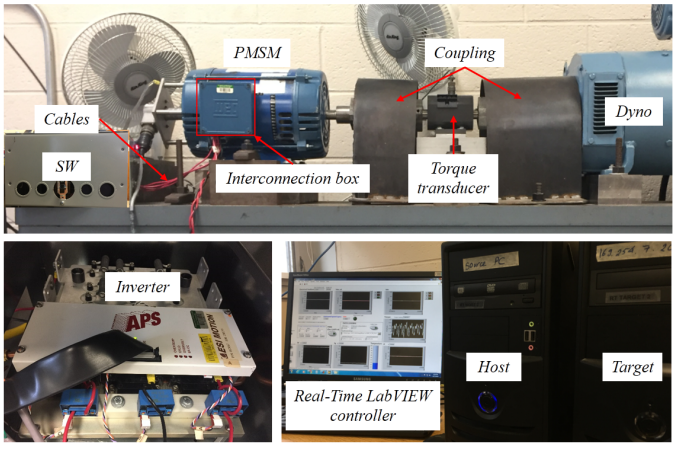


Fig. 6. Experimental setup for ITSCs detection.

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

and the FPGA provides independent 16-bit A/D converter to ensure the high resolution for data sampling.

The machine is operated in constant torque mode at maximum torque per ampere, with decoupling current controller. The angle of current vector was set at $\delta = 110°$ for both co-simulations and experiments. Low controller bandwidth increases the response time of machine drive system. The time constant of PMSM is determined as $\tau = \min\{\frac{L_d}{r_s}, \frac{L_q}{r_s}\}$, and the corresponding bandwidth is $\alpha = \frac{2\pi}{\tau}$. For the PMSM in this work, α is obtained as $304\,\mathrm{rad/s}$ ($L_d = 31.03\,\mathrm{mH}$, $L_q = 62.4\,\mathrm{mH}$). The dynamic startups of the PMSM drive under different current controller bandwidths are shown in Fig. 7. It shows that response time at $200\,\mathrm{rad/s}$ is more than $0.35\,\mathrm{s}$, much longer than that at other bandwidths, and i_d at $200\,\mathrm{rad/s}$ has higher overshoot, which means that controller gains in this case are not good enough. To limit the longest response time, the smallest current controller bandwidth is chosen as $400\,\mathrm{rad/s}$, higher than α. Besides, the available maximum bandwidth depends on the current sampling method and PWM update intervals at a given switch frequency [25]. In both co-simulation and experiment, the maximum bandwidth cannot higher than $2000\,\mathrm{rad/s}$, otherwise the command voltage reaches to the ceiling of dc-link and the system cannot reach steady state.

A. Performance of Co-simulation

Co-simulations provide the ability to investigate harmonics in currents and voltages, which is validated through experiment at the same operating condition and same fault severity. ITSC severity is defined by number of shorted turns and the insulation degradation, and it is not possible to introduce a clean short-circuit experimentally, because the fault can only be introduced by wires and switch outside the machine. Compared to the resistance of each turn, which is about $5\,\mathrm{m\Omega}$, the external resistance is high enough to limit i_f to a low level, less than $10\,\mathrm{A}$, when one turn is shorted. Hence, a fault resistor $r_f = 35\,\mathrm{m\Omega}$, which is close to the external resistance of wires and switch in experiment, is added to the ITSC path in co-simulation to make its incipient fault level the same as that in experiment.

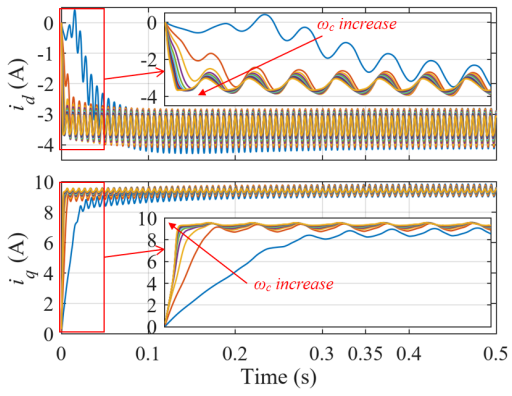


Fig. 7. The dynamic startups of PMSM drive under different current controller bandwidths in co-simulation, ω_c from 200 to $2000\,\mathrm{rad/s}$ with a step of 200, at $10\,\mathrm{A}$ and $300\,\mathrm{rpm}$.

As the bandwidth increases, the harmonics in currents decrease, while those in voltages increase as discussed earlier. The variations of the 6^{th} order harmonics in dq-axis currents and voltages under different bandwidths can be seen in Fig. 8, and it is shown that the currents and voltages in co-simulation have the same behaviors as that in experiment.

Tests are presented as follows: one-turn clean ITSC was detected first because it was only possible in co-simulation (No.1); then the detection of one-turn incipient ITSCs was investigated in both co-simulation (No.2) and experiment (No.3); the same incipient fault was tested in experiment under higher speed (No.4) and higher torque (No.5); finally, the performance of the proposed indicator was tested with more shorted turns in experiment (No.6).

B. One-Turn ITSCs Detection at Low Speed

The proposed indicator was evaluated at $300\,\mathrm{rpm}$ to verify its performance at low speed, as the ITSC is generally difficult to detect at low speed [8]. First, one-turn clean ITSCs in co-simulation were investigated. Then the detection of one-turn incipient ITSCs in both co-simulation and experiment was also investigated.

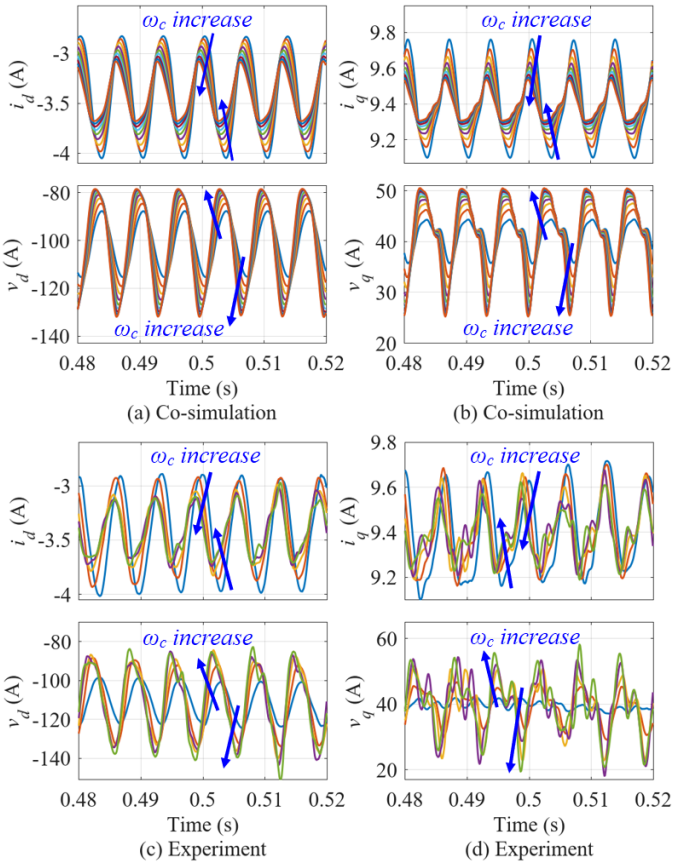


Fig. 8. The dq-axis currents and voltages in one electrical cycle under different ω_c in co-simulation (a) and (b), and in experiment (c) and (d), at $10\,\mathrm{A}$ and $300\,\mathrm{rpm}$. ω_c is from 400 to $2000\,\mathrm{rad/s}$ with a step of 200 for co-simulation and 400 for experiment. Arrows indicate the increase of ω_c. Severity: 1 turn shorted by $r_f = 35\,\mathrm{m\Omega}$ in co-simulation, and directly shorted in experiment. Introducing ITSCs at $0.5\,\mathrm{s}$.

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

1) Clean ITSCs Detection: One-turn clean ITSCs were tested in co-simulation at $I_s = 10\,\text{A}$ and $300\,\text{rpm}$ (No.1), to validate the performance of Ray in fault detection in real applications. Compared to the flux linkage due to the healthy turns in phase A, one shorted turn produces a low opposite flux linkage and has a small influence on the main components in the signals. The dc components and the 6th order harmonics in dq-axis currents and voltages are not affected much by the fault, as shown in Fig. 8. However, the 2nd harmonics increase, which implies that the fault detection using the 2nd harmonics is more sensitive than using the change of dc component when a low proportion of turns are shorted. Fig. 9 shows the amplitudes of the 2nd harmonic in i_{dq} and v_{dq} under different ω_c, including both the healthy and faulty cases, which were marked with H and F, respectively. It shows that $|i_{d2}|$ and $|i_{q2}|$ decrease with the increase of ω_c, and $|v_{d2}|$ and $|v_{q2}|$ increase with the increase of ω_c, which match the trends from analytical model shown in Fig. 3. In the co-simulation of healthy condition, small 2nd harmonic currents and voltages were noticed, due to limited current controller bandwidth, mesh and numerical calculation.

With a minimum value of 654 at $2000\,\text{rad/s}$, Ray ratios are higher and more robust than current and voltage ratios, as shown in Fig. 10. To evaluate the improvement, Ray ratio was compared with the ratio of the 2nd harmonic in IRP after ITSC to that before ITSC (called IRP ratio), presented in test No.1 of Table II. It is shown that Ray ratios are higher than IRP ratios.

2) Incipient ITSCs Detection: The detection of ITSCs with one turn shorted by $r_f = 35\,\text{m}\Omega$ in co-simulation (No.2) and one turn directly shorted in experiment (No.3) were tested at $I_s = 10\,\text{A}$ and $300\,\text{rpm}$. Still, $|i_{d2}|$ and $|i_{q2}|$ decrease with the increase of ω_c, and $|v_{d2}|$ and $|v_{q2}|$ increase with the increase of ω_c, as shown in Fig. 11.

Because of inherent imbalance in experiment, the 2nd harmonic currents and voltages in healthy condition were relatively high, for example, H-$|i_{d2}|$ under $400\,\text{rad/s}$ in experiment was about $30\,\text{mA}$, while it was less than $0.5\,\text{mA}$ in co-simulation, as shown in Fig. 11(c) and (a). Also, the effects of imbalance caused by ITSCs in co-simulation were higher than that in experiment. Therefore, the current and voltage ratios, as well as Ray ratios, in co-simulation were higher than that in experiment, for example, a_{id} and Ray ratio under $400\,\text{rad/s}$ in co-simulation were 4.9 and 17.3, respectively, while that in experiment were 1.38 and 1.65, as shown in Fig. 12(a) and (b). With the minimum values of 15.7 in co-simulation (No.2) and 1.577 in experiment (No.3), both at $2000\,\text{rad/s}$, Ray ratios

TABLE II
RAY RATIOS AND IRP RATIOS

No.		ω_c (100rad/s)								
		4	6	8	10	12	14	16	18	20
1	Ray	710	719	709	696	680	670	660	656	654
	IRP	21.2	22.2	23.7	24.5	25.1	25.3	25.35	25.4	25.5
2	Ray	17.3	17.1	16.6	16.4	16.3	16.1	15.9	15.8	15.7
	IRP	2.51	2.86	3.14	3.32	3.48	3.58	3.70	3.76	3.83
3	Ray	1.65	1.64	1.63	1.61	1.60	1.59	1.588	1.58	1.577
	IRP	1.13	1.16	1.18	1.21	1.22	1.23	1.236	1.239	1.24
4	Ray	1.89	1.86	1.75	1.69	1.67	1.667	1.66	1.652	1.65
	IRP	1.14	1.18	1.19	1.23	1.26	1.28	1.30	1.31	1.32
5	Ray	1.63	1.58	1.56	1.543	1.545	1.54	1.54	1.539	1.538
	IRP	1.16	1.18	1.19	1.20	1.21	1.22	1.23	1.235	1.24
6	Ray	12.0	12.1	12.1	11.8	11.6	11.3	11.2	11.1	11.08
	IRP	2.77	2.83	2.87	3.00	3.06	3.20	3.31	3.346	3.35

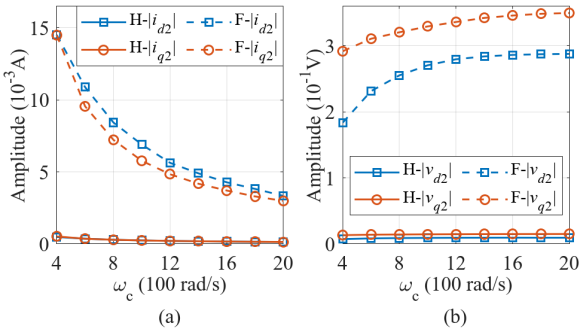


Fig. 9. Amplitudes of the 2nd harmonics in dq-axis currents (a) voltages (b) under different ω_c, at 10 A and 300 rpm in co-simulation. Severity: 1 turn shorted by r_f = 0. Healthy and faulty cases are marked with H and F, respectively.

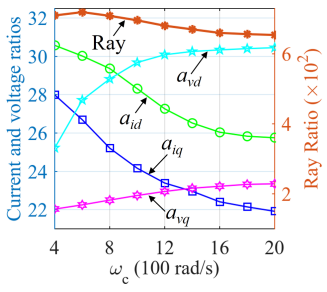


Fig. 10. Current and voltage ratios and Ray ratios under different ω_c, at 10 A and 300 rpm in co-simulation. Severity: 1 turn shorted by r_f = 0.

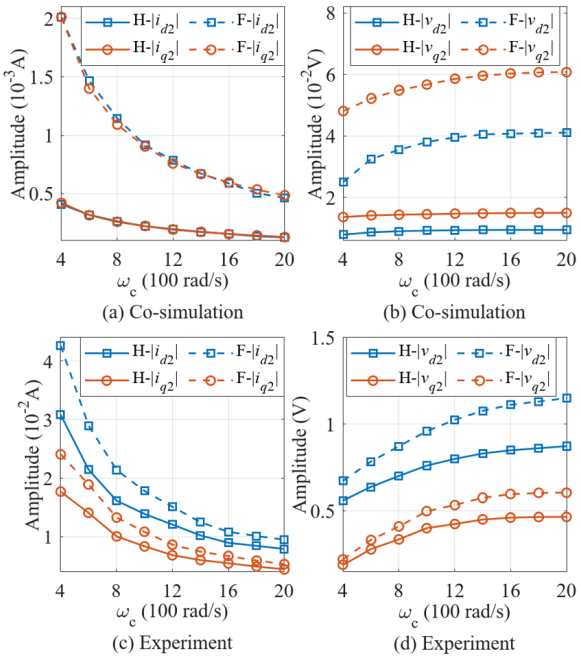


Fig. 11. Amplitudes of the 2nd harmonics in dq-axis currents and voltages under different ω_c, at 10 A and 300 rpm. Severity: 1 turn shorted by r_f = 35 mΩ in co-simulation, (a) and (b), and directly shorted in experiment, (c) and (d).

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

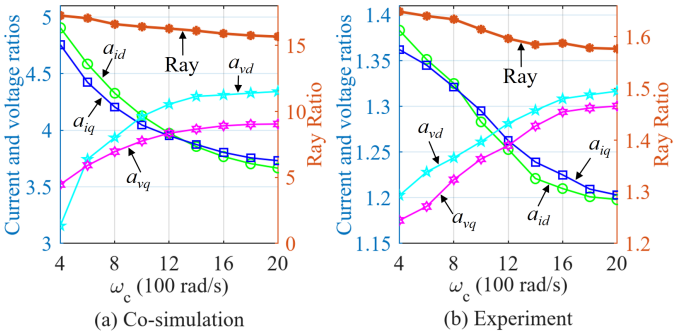


Fig. 12. Current and voltage ratios and Ray ratios under different ω_c, at 10 A and 300 rpm. Severity: 1 turn shorted by r_f = 35 mΩ in co-simulation (a), and directly shorted in experiment (b).

were higher and more robust than current and voltage ratios, shown in Fig. 12(a) and (b), respectively. Ray ratios are also higher than IRP ratios, as presented in tests No.2 and No.3 of Table II.

Ray detects ITSCs faster than IRP due to its higher indicator value, as shown in Fig. 13, where ITSCs were introduced at 0.5 s and the threshold was set as 1.2 to eliminate the false alarm caused by the noise in the experiment. Under a wide bandwidth range, ω_c from 400 to 2000 rad/s, the detection time by Ray was between 59 ms to 87 ms, and that by IRP was between 115 ms to 156 ms. The faults were undetectable by IRP at 400 rad/s (IRP-ω_c400) and 800 rad/s (IRP-ω_c800), which validates that Ray is more reliable for ITSCs detection.

C. One-Turn ITSCs Detection at High Speed and High Torque

In experiment, the proposed indicator was tested at higher speed, 600 rpm with I_s = 10 A (No.4), and higher commanded toque, I_s = 20 A with 300 rpm (No.5), for the same severity in Section V-B, one turn directly shorted. With the minimum values of 1.65 and 1.538 for test No.4 and No.5, respectively, both at 2000 rad/s, the Ray ratios were higher than the current and voltage ratios, as shown in Fig. 14(a) and (c). Introducing ITSCs at 0.5 s under different ω_c, the faults were detected within 25 ms to 44 ms in test No.4, and 64 ms to 82 ms in test No.5, both within 2.5 electrical cycles (50 ms and 100 ms for 600 rpm and 300 rpm, respectively), as shown in Fig. 14(b) and (d). Ray ratios are also higher than IRP ratios, as presented in tests No.4 and No.5 of Table II.

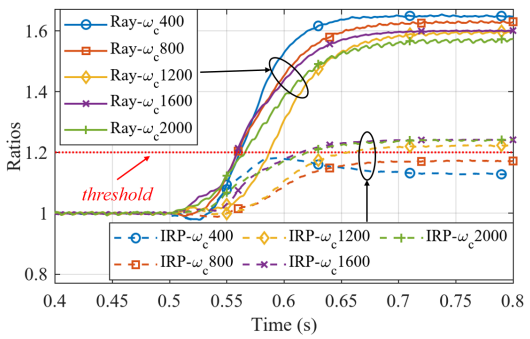


Fig. 13. Ray ratios and IRP ratios for online ITSCs detection under different ω_c, at 10 A and 300 rpm in experiment. Severity: 1 turn directly shorted.

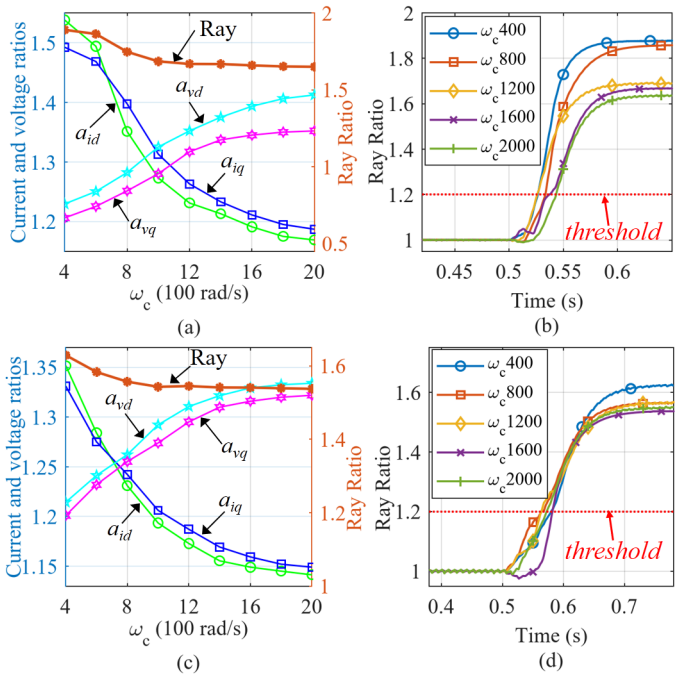


Fig. 14. Current and voltage ratios and Ray ratios (a) and (c), and online ITSCs detection (b) and (d) under different ω_c, at 600 rpm with 10 A (a) and (b), and 20 A with 300 rpm (c) and (d) in experiment. Severity: 1 turn directly shorted.

D. Fifteen-Turn ITSCs Detection

To further investigate the effectiveness of Ray, experiments were conducted at I_s = 10 A and 600 rpm for ITSCs with 15 turns, and a fault resistor r_f = 0.5 Ω was added to limit i_f and avoid the machine failure (No.6). The current and voltage ratios of test No.6 are between 2.5 to 4.7, as shown in Fig. 15(a), which were higher than the cases of one-turn incipient ITSC (No.3 to No.5) because of the larger machine imbalance caused by the 15 shorted turns. With a minimum value of 11.08 at 2000 rad/s, Ray ratios were higher and more robust than current and voltage ratios. The ITSCs could be detected within 9 ms to 15 ms, as shown in Fig. 15(b), faster than the cases of one-turn incipient ITSC due to the high indicator values. Again, Ray ratios are also higher than IRP ratios, as presented in tests No.6 of Table II.

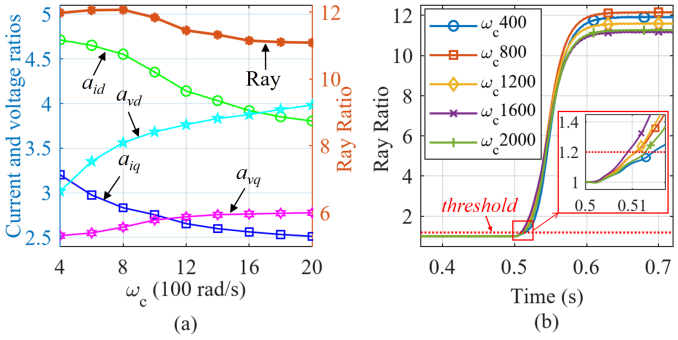


Fig. 15. Current and voltage ratios and Ray ratios (a) and online ITSCs detection under different ω_c (b), at 10 A with 600 rpm in experiment. Severity: 15 turn shorted by r_f = 0.5 Ω.

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

E. Discussion

1) Circulating Current: i_f contains mainly the fundamental, 3rd, 5th and 7th harmonics. As ω_c increase, the fundamental is almost not affected, while the 3rd, 5th and 7th harmonics increase, due to the increasing corresponding harmonics in EMF voltages, as shown in Fig. 16. The rms value of i_f in each test also increase slightly, as shown in Table III. The circulating current is fed by two quantities: the EMF of faulted turns, e_{af}, and the faulty phase current, i_a [10]. Thus, the rms values of i_f at 600 rpm (No.4) were higher than that at 300 rpm (No.3) because the EMF is proportional to the speed, and the rms values at 20 A (No.5) were also higher than that at 10 A (No.3) because more contribution from i_a.

It was demonstrated that ITSC with one shorted turn leads to the highest fault current [8], which is, however, only valid for clean ITSC (No.1). For incipient fault, the worst scenario is no longer the single turn fault [10], as shown by the comparison of one-turn incipient fault (No.4) and fifteen-turn incipient fault (No.6) in Table III, where the rms values of i_f are nearly the same.

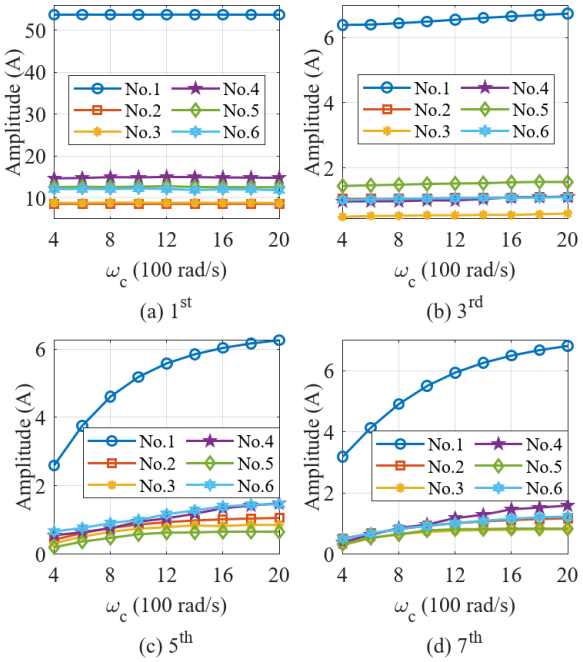


Fig. 16. The harmonics in i_f under different ω_c. From (a) to (d): the 1st, 3rd, 5th and 7th harmonic components.

TABLE III
THE RMS OF CIRCULATING CURRENT (A)

No.	N_f	r_f (mΩ)	I_s (A)	ω_m (rpm)	ω_c (100rad/s) 4	8	12	16	20
1	1	0	10	300	38.4	38.7	38.8	38.9	38.92
2	1	35	10	300	6.14	6.19	6.21	6.23	6.24
3	1	0 [a]	10	300	6.34	6.48	6.50	6.73	7.37
4	1	0 [a]	10	600	9.94	10.5	10.9	11	11.3
5	1	0 [a]	20	300	8.75	9.06	9.16	9.24	9.31
6	15	500 [a]	10	600	9.95	10.5	10.7	11	11.22

[a] In experiment, r_f = 0 and 500 mean the directly shorted and ITSC with an additional 0.5 Ω resistor, respectively.

2) Indicator value: Since higher speed induces higher back EMF and hence higher circulating current, the 2nd harmonic currents and voltages also increase. Hence, Ray ratios under different ω_c at 600 rpm (No.4) were higher than that under the corresponding ω_c at 300 rpm (No.3), as presented in Table II. While, doubling the commanded current from 10 A to 20 A does not increase Ray ratio. This is because more negative i_d is injected to the machine, which reduces the effect of magnet on the fault information in currents and voltages. Therefore, the indicator values under different ω_c at 20 A (No.5) were not higher than that under the corresponding ω_c at 10 A (No.3), which can also be found in Table II.

To study the effect of torque on indicator value, experiments were performed with one turn directly shorted from 2.5 A to 25 A with a step of 2.5 A. The speed was fixed at 300 rpm and the current controller bandwidth was fixed at ω_c = 400 rad/s. It shows that $|i_{d2}|$, $|i_{q2}|$ and $|v_{d2}|$ increase with the increase of I_s, and $|v_{q2}|$ decreases first at low I_s but increase at high I_s, as shown in Fig. 17. As a result, Ray ratio peaks in 15 A, as shown in Table IV. Both under 400 rad/s, the variation of Ray ratio within 2.5 A to 25 A at 300 rpm is from 1.51 to 1.72, and that from 300 rpm (No.3) to 600 rpm (No.4) at 10 A is from 1.65 to 1.89, which means that the effect of speed on fault detection is greater than of torque.

3) Robustness: Coefficient of variation, CV, is used to evaluate the robustness with respect to ω_c of the proposed ITSC indicator.

$$\mathrm{CV} = \frac{\sigma_x}{\mu_x} \tag{21}$$

where σ_x and μ_x are the standard deviation and the mean of a_{id}, a_{iq}, a_{vd}, a_{vq}, Ray ratio and IRP ratio, respectively. Hence, small CV means high robustness. The CVs of tests No.1 to No.6 are listed in Table V. It shows that the CV of Ray ratio in each test is smaller than most of the corresponding current and voltage ratios, and hence Ray ratio is more robust than

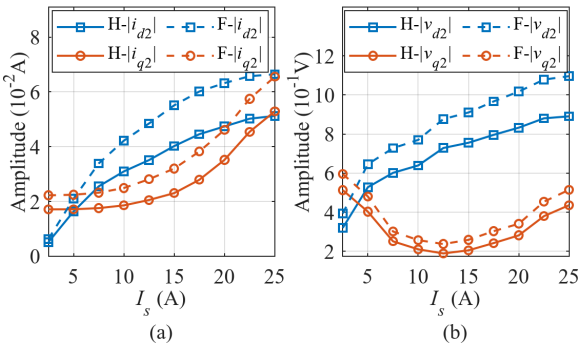


Fig. 17. Amplitudes of the 2nd harmonics in dq-axis currents (a) and voltages (b) under different I_s, at 300 rpm in experiment. Severity: 1 turn directly shorted.

TABLE IV
CURRENT AND VOLTAGE RATIOS AND RAY RATIOS AT DIFFERENT I_s

I_s(A)	2.5	5	7.5	10	12.5	15	17.5	20	22.5	25
i_d	1.22	1.29	1.33	1.36	1.38	1.37	1.35	1.34	1.31	1.3
i_q	1.3	1.31	1.32	1.34	1.37	1.39	1.37	1.31	1.27	1.24
v_d	1.23	1.22	1.21	1.205	1.203	1.21	1.215	1.22	1.225	1.23
v_q	1.16	1.19	1.2	1.23	1.25	1.26	1.25	1.21	1.2	1.18
Ray	1.51	1.57	1.61	1.65	1.7	1.72	1.69	1.62	1.56	1.53

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

TABLE V
THE CVS OF TESTS NO.1 TO NO.6 ($\times 10^{-2}$)

No.	a_{id}	a_{iq}	a_{vd}	a_{vq}	Ray ratio	IRP ratio
1	6.97	8.84	6	2.91	3.52	6.49
2	10.44	8.48	9.7	4.8	3.42	13.18
3	5.49	4.68	3.2	3.77	1.75	3.26
4	10.74	9.06	5.23	4.79	4.92	5.16
5	6.08	5.22	3.4	3.38	1.96	2.42
6	8.36	8.48	8.54	3.75	3.56	7.48

most of the current and voltage ratios. Compared to the CVs of IRP ratio in each test, the CVs of Ray ratio are also smaller. Therefore, the ITSCs detection by Ray ratio is more robust with respect to current controller bandwidth.

4) Detection Time: The detection time for one-turn ITSCs and fifteen-turn ITSCs in experiment under different operating conditions are listed in Table VI. It shows that the fault detection at higher speed becomes faster because the harmonic extracting technique is adaptive to machine speed. As the speed increase, the periods of currents and voltages decreases, and hence the time available to extracting the 2nd harmonics decreases, as the comparison of tests No.3 and No.4 in Table VI. There is no big difference between the fault detection at 20 A (No.5) and 10 A (No.3), mainly because the currents and voltages in these two operating conditions have the same period and the Ray ratios are approximately the same.

Overall, the proposed indicator can detect one-turn faults under different ω_c within 2.5 electrical cycles, and the detection of fifteen-turn faults (No.6) is much faster due to the higher indicator values.

VI. CONCLUSION

In this paper, a novel indicator based on the fault information in both currents and voltages is proposed for ITSCs detection. The faults change the dc components of i_{dq} and v_{dq} and also introduce the 2nd order harmonics. For a closed current-loop controlled PMSM with a specific fault severity, the 2nd harmonics are mainly affected by the current controller bandwidth.

The mathematical models of the 2nd harmonic currents and voltages were derived, and theoretical analysis, simulations and experiments showed that the 2nd harmonics in i_{dq} decrease as the bandwidth increases, and that the trends in v_{dq} were opposite. To combine the 2nd harmonics in currents and voltages, a new function, Rayleigh quotient, was introduced to weight the current and voltage ratios under different bandwidths, and the resulting Ray ratio was evaluated as the ITSCs indicator. The robustness of Ray ratio with respect to current controller bandwidth was validated by faults detection in co-simulation and in experiment.

TABLE VI
THE FAULTS DETECTION TIME IN EXPERIMENT (ms)

No.	N_f	r_f (mΩ)	I_s (A)	ω_m (rpm)	ω_c (100rad/s) 4	8	12	16	20
3	1	0	10	300	61.3	59.4	87	58.6	66
4	1	0	10	600	26.4	33.4	25.8	37.1	43.3
5	1	0	20	300	79.6	64	63.5	81.6	72.2
6	15	500	10	600	14.5	10.8	10.6	9.1	12.8

The detection of ITSCs with one-turn (i.e., less than 0.4% of the phase winding) under different operating conditions, and that with fifteen-turn (i.e., 5% of the phase winding) at low speed proved the effectiveness of the proposed indicator. It also showed that the proposed indicator has a more robust and faster detection performance than IRP ratio.

The circulating current and the indicator value increase as the speed increase, due to the higher back EMF. The fault detection at higher speed becomes faster, because the duration of extracting the harmonics in currents and voltages decrease. While, the indicator value is not proportional to the torque, because the high negative d-axis current demagnetizes the magnet at higher torque. And hence, the fault detection at higher torque is not always faster than that at lower torque. For ITSC with more shorted turns, the fault becomes easier to detect with higher indicator value and shorter detection time.

REFERENCES

[1] M. Drif and A. J. M. Cardoso, "Stator fault diagnostics in squirrel cage three-phase induction motor drives using the instantaneous active and reactive power signature analyses," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1348–1360, May 2014.

[2] A. Gandhi, T. Corrigan, and L. Parsa, "Recent advances in modeling and online detection of stator interturn faults in electrical motors," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1564–1575, May 2011.

[3] S. N. Foster, J. G. Cintron-Rivera, and E. G. Strangas, "Detection of incipient stator winding faults in PMSMs with single-layer fractional slot concentrated windings," *Electric Power Systems Research*, vol. 131, pp. 231–243, Feb 2016.

[4] H. H. Eldeeb, A. Berzoy, and O. Mohammed, "Stator fault detection on DTC-driven IM via magnetic signatures aided by 2-D FEA co-simulation," *IEEE Transactions on Magnetics*, vol. 55, no. 6, pp. 1–5, June 2019.

[5] T. Boileau, N. Leboeuf, B. Nahid-Mobarakeh, and F. Meibody-Tabar, "Synchronous demodulation of control voltages for stator interturn fault detection in PMSM," *IEEE Transactions on Power Electronics*, vol. 28, no. 12, pp. 5647–5654, Dec 2013.

[6] J. Urresty, J. Riba, and L. Romeral, "Diagnosis of interturn faults in PMSMs operating under nonstationary conditions by applying order tracking filtering," *IEEE Transactions on Power Electronics*, vol. 28, no. 1, pp. 507–515, Jan 2013.

[7] S. Huang, E. G. Strangas, A. Aggarwal, K. Li, and F. Niu, "Robust interturn short-circuit detection in PMSMs with respect to current controller bandwidth," in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sep. 2019, pp. 3897–3904.

[8] B. Wang, J. Wang, A. Griffo, and B. Sen, "Stator turn fault detection by second harmonic in instantaneous power for a triple-redundant fault-tolerant PM drive," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 7279–7289, Sept 2018.

[9] B. M. Ebrahimi and J. Faiz, "Feature extraction for short-circuit fault detection in permanent-magnet synchronous motors using stator-current monitoring," *IEEE Transactions on Power Electronics*, vol. 25, no. 10, pp. 2673–2682, Oct 2010.

[10] J. G. Cintron-Rivera, S. N. Foster, and E. G. Strangas, "Mitigation of turn-to-turn faults in fault tolerant permanent magnet synchronous motors," *IEEE Transactions on Energy Conversion*, vol. 30, no. 2, pp. 465–475, June 2015.

[11] R. Z. Haddad, C. A. Lopez, S. N. Foster, and E. G. Strangas, "A voltage-based approach for fault detection and separation in permanent magnet synchronous machines," *IEEE Transactions on Industry Applications*, vol. 53, no. 6, pp. 5305–5314, Nov 2017.

[12] G. Ahn, J. Lee, C. H. Park, M. Youn, and B. D. Youn, "Inter-turn short circuit fault detection in permanent magnet synchronous motors based on reference voltage," in *2019 IEEE 12th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, Aug 2019, pp. 245–250.

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPEL.2020.3030036, IEEE Transactions on Power Electronics

[13] H. Saavedra, J.-C. Urresty, J.-R. Riba, and L. Romeral, "Detection of interturn faults in PMSMs with different winding configurations," *Energy Conversion and Management*, vol. 79, pp. 534–542, 03 2014.

[14] J. Faiz, A. H. Exiri, and H. Nejadi-Koti, "Current-based inter-turn short circuit fault modeling in permanent magnet synchronous machine using magnetic equivalent circuit model," in *2016 10th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, June 2016, pp. 265–270.

[15] R. Z. Haddad and E. G. Strangas, "On the accuracy of fault detection and separation in permanent magnet synchronous machines using MCSA/MVSA and LDA," *IEEE Transactions on Energy Conversion*, vol. 31, no. 3, pp. 924–934, Sep. 2016.

[16] W. L. Roux, R. G. Harley, and T. G. Habetler, "Detecting faults in rotors of PM drives," *IEEE Industry Applications Magazine*, vol. 14, no. 2, pp. 23–31, March 2008.

[17] M. Eltabach, A. Charara, and I. Zein, "A comparison of external and internal methods of signal spectral analysis for broken rotor bars detection in induction motors," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 1, pp. 107–121, Feb 2004.

[18] C. Kral, R. S. Wieser, F. Pirker, and M. Schagginger, "Sequences of field-oriented control for the detection of faulty rotor bars in induction machines-the vienna monitoring method," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 1042–1050, Oct 2000.

[19] M. B. Abadi, S. M. A. Cruz, A. P. Gonçalves, P. F. C. Gonçalves, and A. M. S. M. *et al.*, "Detection of stator and rotor faults in a DFIG based on the stator reactive power analysis," in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, Oct 2014, pp. 2037–2043.

[20] B. Vaseghi, B. Nahid-Mobarakeh, N. Takorabet, and F. Meibody-Tabar, "Experimentally validated dynamic fault model for PMSM with stator winding inter-turn fault," in *2008 IEEE Industry Applications Society Annual Meeting*, Oct 2008, pp. 1–5.

[21] B. Du, S. Wu, S. Han, and S. Cui, "Interturn fault diagnosis strategy for interior permanent-magnet synchronous motor of electric vehicles based on digital signal processor," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 3, pp. 1694–1706, March 2016.

[22] L. Romeral, J. C. Urresty, J. R. R. Ruiz, and A. G. Espinosa, "Modeling of surface-mounted permanent magnet synchronous motors with stator winding interturn faults," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1576–1585, May 2011.

[23] M. Zafarani, E. Bostanci, Y. Qi, T. Goktas, and B. Akin, "Interturn short-circuit faults in permanent magnet synchronous machines: An extended review and comprehensive analysis," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 6, no. 4, pp. 2173–2191, Dec 2018.

[24] Y. Qi, E. Bostanci, V. Gurusamy, and B. Akin, "A comprehensive analysis of short-circuit current behavior in PMSM interturn short-circuit faults," *IEEE Transactions on Power Electronics*, vol. 33, no. 12, pp. 10 784–10 793, Dec 2018.

[25] S.-K. Sul, *Control of Electric Machine Drive Systems*, ser. IEEE Press Series on Power Engineering. Wiley, 2011.

Shaopo Huang received the B.S. and M.S. degrees from Hebei University of Technology (HEBUT), Tianjin, China, in 2009 and 2012, respectively, both in electrical engineering. He is currently working toward the Ph.D. degree with the School of Electrical Engineering, HEBUT.

He was an Engineer with ABB LV Installation Materials Co., Ltd. Beijing, China, from 2012 to 2015. From 2017 to 2019, he was a Research Fellow with the Electrical Machines and Drives Laboratory, Michigan State University, East Lansing, MI, USA. His research interests include fault diagnosis and mitigation of electrical drive systems, and reliability testing of electrical apparatus.

Anmol Aggarwal received the B.E. degree in electrical engineering from the Delhi Technological University, Delhi, India, in 2013. He is currently working towards the Ph.D. degree in the Electrical Machines and Drives Laboratory, Michigan State University, East Lansing, MI, USA.

His research interests include design of segmented stators for permanent magnet synchronous machines, control and fault diagnosis of electrical machines and drives.

Elias G. Strangas (M'80-SM'15) received the Dipl. Eng. degree from the National Technical University of Greece, Athens, Greece, in 1975, and the Ph.D. degree from the University of Pittsburgh, Pittsburgh, PA, USA, in 1980, both in electrical engineering.

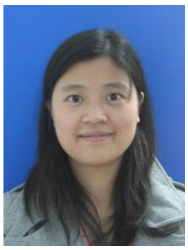
He was with Schneider Electric, Athens, from 1981 to 1983, and with the University of Missouri, Rolla, MO, USA, from 1983 to 1986. Since 1986, he has been with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI, USA, where he heads the Electrical Machines and Drives Laboratory. His research interests include the design and control of electrical machines and drives, finite-element methods for electromagnetics, and failure prognosis and mitigation of electrical drive systems.

Kui Li was born in Baoding, China, in 1965. He received the B.S. and M.S. degrees in electrical engineering from the Hebei University of Technology, Tianjin, China, in 1987 and 1992, respectively, and the Ph.D. degree in electrical engineering from Fuzhou University, Fuzhou, China, in 1996.

He is currently a Professor with the School of Electrical Engineering, Hebei University of Technology. He has authored or coauthored over 100 technical articles and two monographs. His current research interests include reliability and intellectualization of electrical apparatus, fault diagnosis, and life prediction of electrical apparatus.

Feng Niu (M'15) was born in Hebei, China, in 1986. He received the B.S and Ph.D. degrees from the Hebei University of Technology, Tianjin, China, in 2009 and 2015, respectively, both in electrical engineering.

From 2012 to 2014, he was a Research Fellow with the Electrical Machines and Drives Laboratory, Michigan State University, East Lansing, MI, USA. From 2016 to 2018, he was a Post-Doctoral Research Fellow with the College of Electrical Engineering, Zhejiang University, Hangzhou, China. He is currently a Professor with the School of Electrical Engineering, Hebei University of Technology. He has authored or coauthored over 40 technical articles. His current research interests include motor system and control, and intelligent electrical equipment.



Xiaoyan Huang (M'09) received the B.E. degree in control measurement techniques and instrumentation from Zhejiang University, Hangzhou, China, in 2003, and received the Ph.D. degree in electrical machines and drives from the University of Nottingham, Nottingham, U.K., in 2008.

From 2008 to 2009, she was a Research Fellow with the University of Nottingham. She is currently a Professor with the College of Electrical Engineering, Zhejiang University, Hangzhou, China, where she is working on electrical machines and drives. Her research interests include PM machines and drives for aerospace and traction applications, and generator systems for urban networks.

0885-8993 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 05,2020 at 05:22:04 UTC from IEEE Xplore. Restrictions apply.