SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(65th Edition)

MAY, 2018

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 **Sixty 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 **May, 2018**. 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. A2B corroles: Fluorescence signaling systems for sensing fluoride ions, ***6.Om prakash Yadav***, ***6.Atul Varshney***, ***3.Anil Kumar***, ***6.Ratneshwar Kumar Ratnesh*** and ***3.Mohan Singh Mehata***, Applied Chemistry and Applied Physics, DTU

2. A Comprehensive Study of Text Summarization Algorithms, Yash Dhankhar, ***7.Indu Bala***, ***6.Swati Singh*** and Sunil Dalal, IT Computer, DTU

3. A Deep Structure of Person Re-Identification using Multi-Level Gaussian Models, ***3.Dinesh Kumar Vishwakarma*** and ***7.Sakshi Upadhyay***, ECE, DTU

4. A Hand-Based Biometric Verification System Using Ant Colony Optimization, ***6.Nirmal Pandey***, ***3.O.P. Verma*** and ***6.Amioy Kumar***, Computer Science, DTU

5. A novel fuzzy energy based level set method for medical image segmentation, ***3.Mahipal Singh Choudhry*** and ***3.Rajiv Kapoor***, Electronics & Communication, DTU

6. A STUDY ON INCREMENTAL MINING OF FREQUENT PATTERNS, M. Dutta Borah and ***3.R. Jindal***, Computer Science, DTU

7. Assessment of Academic Departments Efficiency Using Data Envelopment Analysis, Sushil Yadav*, Lakshay, ***8.Chitransh Joshi***, Mandeep Singh, ***8.Gurmeet Singh Bhatia*** , ***3.Ranganath M. Singari*** and ***3.Nand Kumar***, Mechanical, DTU

8. Boosting performance of power quality event identification with KL Divergence measure and standard deviation, ***3.Rajiv Kapoor***, Rashmi Gupta, Le Hoang Son, Sudan Jha and Raghvendra Kumar, ECE, DTU

9. Cement Industry and Strategies for Mitigating Carbon Emissions, ***6.Ahana Ghosh*** and ***3.Anubha Mandal***, Environmental, DTU

10. Cheater Detection Scheme for Dynamic Secret Key Generation, Sunil Dalal, ***6.Sangeeta Bhan*** and Susobhan Das, Computer Science, DTU

11. Design of Flying Wing UAV and Effect of Winglets on its Performance, ***8.Rishabh Dagur***, ***8.Vikrant Singh***, ***8.Shabir Grover***, ***8.Nikhil Sethi***, ***3.B.B. Arora***, Mech, DTU

12. Designing of Supplementary Controller for STATCOM for Damping of Oscillations in Power System, Vipin Jain and ***3.Narendra Kumar***, Electrical, DTU

13. Digital Video Watermarking: Issues and Challenges, Md Shahid and ***7.1.Pradeep Kumar***, Computer, DTU

14. DroidMark – A Tool for Android Malware Detection using Taint Analysis and Bayesian Network, ***6.Dhruv Rathi*** and ***3.Dr Rajni Jindal***, Computer, DTU

15. Dy3+ ions doped oxy-fluoro boro tellurite glasses for the prospective optoelectronic device applications, V. Murali Krishna, Sk. Mahamuda, Rupesh A. Talewar, K. Swapna, M. Venkateswarlu and ***3.A. S. Rao***, Physics, DTU

16. Dynamical Model of Epidemic Along with Time Delay; Holling Type II Incidence Rate and Monod–Haldane Type Treatment Rate, ***6.Abhishek Kumar*** and ***3.Nilam***, Applied Mathematics, DTU

17. Effect of Non-linearity on Wheel/rail Interaction Dynamics using Bond Graph, Vivek Kumar, ***3.Vikas Rastogi*** and P.M. Pathak, Mechanical, Production, Industrial and Automobile, DTU

18. Effective Features to Classify Big Data using Social Internet of Things, Lakshmanaprabu SK, Shankar K, Ashish Khanna, ***7.1.Deepak Gupta***, Joel J. P. C. Rodrigues, Plácido R. Pinheiro, Victor Hugo C. de Albuquerque, CSE, DTU

19. Effective methods of waste management in India, ***3.R.S. Mishra***, ***3.P.V Ram Kumar***, ***3.Amit Pal***, ***6.Mayank Kumar Singh***, Mechanical & Production, DTU

20. Electrochemical genosensor based on template assisted synthesized polyaniline nanotubes for chronic myelogenous leukemia detection, Amrita Soni, ***6.Chandra Mouli Pandey***, Shipra Solanki, R.K. Kotnala and Gajjala Sumana, Applied Chemistry, DTU

21. Electron irradiation induced wall-to-wall joining of multiwalled carbon nanotubes, ***6.Reetu Kumari***, ***3.Pawan K. Tyagi***, ***3.Nitin K. Puri***, Applied Physics, DTU

22. Experimental comparative study of chrome steel pin with and without chrome plated cast iron disc in situ fully flooded interface lubrication, ***3.Roop Lal*** and ***3.R. C. Singh***, Mechanical, DTU

23. Experimental Validation of PID and LQR Control Techniques for Stabilization of Cart Inverted Pendulum System, Sudarshan K.Valluru, Madhusudan Singh, Mayank Singh and Vanshaj Khattar, Electrical, DTU

24. Few-layer p-type phosphorene sheet: An efficient transparent conducting electrode in silicon heterojunction solar cell, Chandra Kamal Borah, ***3.Pawan K. Tyagi***, Sanjeev Kumar and Kamlesh Patel, Applied Physics, DTU

25. Geopolymer concrete: A way of sustainable construction, ***6.Manvendra Verma***, & ***3.Dr Nirendra Dev***, Civil, DTU

26. GroEL Mediates Folding of Bacillus anthracis Serine/Threonine Protein Kinase, PrkC, ***6.Richa Virmani***, Yogendra Singh and ***3.Yasha Hasija***, Biotechnology, DTU

27. ILookLikeAnEngineer: Using Social Media Based Hashtag Activism Campaigns as a Lens to Better Understand Engineering Diversity Issues, Aqdas Malik, ***8.1.Aditya***

28. ***Johri***, RAJAT HANDA, Habib Karbasian and Hemant Purohit, Mechanical, DTU

29. Improved diagnosis of Parkinson's disease using optimized crow search algorithm, ***7.1.Deepak Gupta***, Shirsh Sundaram, Ashish Khanna, Aboul Ella Hassanien, Victor Hugo C. de Albuquerque, Computer science, DTU

30. Mechanism of Joint Formation and Characteristics of Interface in Ultrasonic welding: Literature Review, Bharat Sanga, ***3.Reeta Wattal*** and ***3.D. S. Nagesh***, Mech, DTU

31. Modelling Facebook usage for collaborative learning in higher education, ***6.Anurag Tiruwa***, ***3.Rajan Yadav*** and ***3.Pradeep Kumar Suri***, Delhi School of Management, DTU

32. New scheme for underwater acoustically wireless transmission using direct sequence code division multiple access in MIMO systems, ***3.Rajiv Kapoor***, Rashmi Gupta, Raghavendra Kumar, Le Hoang Son and Sudan Jha, ECE, DTU

33. Optimization of Transfer Learning for Sign Language Recognition Targeting Mobile Platform, ***6.Dhruv Rathi***, Computer Engineering, DTU

34. Performance Enhancement Analysis of Double Effect Vapor Absorption System Using Loop Heat Pipes, ***7.Ankit Dwivedi*** and ***3.R S Mishra***, Mechanical, DTU

35. Photoluminescence investigations on Dy3+ ions doped Zinc Lead Tungsten Tellurite glasses for optoelectronic devices, ***6.Ritu Sharma*** and ***3.A.S. Rao***, A Physics, DTU

36. Product Geometry Analysis of Cyclic Imidates using Molecular Mechanics, ***6.Dhirendra Brahmchari*** and ***3.Saurabh Mehta***, Applied Chemistry, DTU

37. Proposal for Dimensionality Testing in Quantum Private Query, Arpita Maitra, Bibhas Adhikari and ***6.Satyabrata Adhikari***, Applied Mathematics, DTU

38. Protective polyurea coatings for enhanced blast survivability of concrete, ***6.N. Iqbal***, P.K. Sharma, ***3.D. Kumar*** and P.K. Roy, Chemistry and Polymer Technology, DTU

39. Review of Heterocyclic Scaffolds for the Inhibitors of ATP synthase, ***3.Saurabh Mehta***, Applied Chemistry, DTU

40. Review of the Progress for the Synthesis of N- and OHeterocycles through Iodine-mediated Cyclizations, ***Saurabh Mehta***, Applied Chemistry, DTU

41. Review on Effects of Different Parmeters on Behavior of Geopolymer Concrete, ***6.Manvendra Verma*** and ***3.Nirendra Dev***, Civil. DTU

42. Segmentation of Dark Foreground Objects by Maximum Non-Extensive Entropy Partitioning, ***3.Seba Susan***, ***8.Rahul Singh***, ***8.Amit Kumar***, ***8.Abhishek Kumar*** and ***8.Ashwani Kumar***, IT, DTU

43. Sensitization of Nd3+ by 4f-5d transition of Ce3+ in Ba2Y(BO3)2Cl Phosphor for the Prospective NIR Applications, R.A. Talewar, Sk. Mahamuda, ***3.A.S. Rao***, C.P. Joshi, S.V. Moharil, Applied Physics, DTU

44. Simulation of Inductively Coupled Plasma of Ar/C2H2/CH4/H2 gas mixture in PECVD reactor and calculating the reactor efficiency, ***7.Manish Kumar***, ***7.Sagar Khanna*** and ***3.Suresh C Sharma***, Applied Physics, DTU

45. Solar Energy a Path to India's Prosperity, Yogender Pal Chandra, Arashdeep Singh, Vikas Kannojiya and ***3.J. P. Kesari***, Mechanical, DTU

46. Solar Thermal Technology: Potential and Industrial Applications, ***7.Piyush Rawat*** and ***3.J. P. Kesari***, Mechanical, DTU

47. Status of Ambient Air Quality in Khondbond Iron Ore Mines, Orissa (India), Ankita Srivastava, ***6.Amit Krishan*** and ***3.Rajeev Kumar Mishra***, Environmental, DTU

48. Study of Improvement in the Performance of Triple Effect Vapor Absorption System Using Loop Heat Pipes, ***7.1.Ankit Dwivedi*** and ***3.R S Mishra***, Mechanical, DTU

49. Study of Performance Enhancement of Half Effect Vapor Absorption System Using Loop Heat Pipes, ***7.1.Ankit Dwivedi*** and ***3.R S Mishra,*** Mechanical, DTU

50. Subspace-Based Adaptation of Detectors for Video, ***7.Sangeeta Bhan***, Sunil Dalal and Pankaj Choudhary, Computer Science, DTU

51. Surface Composite Fabrication through TIG arc Process: A Review, ***3.N. Yuvaraj***, Mechanical, DTU

52. Synthesis and Characterization of Thiosemicarbazone Metal Complexes, ***7.Shankar Suman*** and ***3.Ram Singh***, Applied Chemistry, DTU

53. Synthetic Strategies for Isochromenes: Recent Advances, ***3.Saurabh Mehta***, Applied Chemistry, DTU

54. Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning, ***6.Dhruv Rathi***, ***Sushant Jain*** and ***3.Dr. S. Indu***, Computer Engineering and Electronics and Communication, DTU

1. *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*

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 202 (2018) 207–213

Contents lists available at ScienceDirect

Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy

journal homepage: www.elsevier.com/locate/saa

A_2B corroles: Fluorescence signaling systems for sensing fluoride ions

Omprakash Yadav [a], Atul Varshney [a], Anil Kumar [a,*], Ratneshwar Kumar Ratnesh [b], Mohan Singh Mehata [b]

[a] Department of Applied Chemistry, Delhi Technological University, Delhi 42, India
[b] Department of Applied Physics, Delhi Technological University, Delhi 42, India

ARTICLE INFO

Article history:
Received 23 September 2017
Received in revised form 10 May 2018
Accepted 14 May 2018
Available online xxxx

Keywords:
Corroles
Fluorescence
Sensor
Fluoride
Tetrapyrrole

ABSTRACT

Four free base corroles, **1**–**4**, A_2B, (where A = nitrophenyl, and B = pentafluorophenyl, 2, 6-difluoro, 3, 4, 5-trifluoro and 4-carboxymethylphenyl group) have been synthesized, characterized and demonstrated as excellent chemosensor for the detection of fluoride ions selectively in toluene solution. The reported corroles shows highest quantum yield in free base form of porphyrinoid systems so far. All these corrole, **1**–**4**, have the excellent ability to sense fluoride ion. Cumulative effect of static and dynamic factors is responsible for the quenching of fluorescence which indicates the detection of fluoride ion in solution.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The established literature reveals that different fluorescence materials have been extensively explored in the domain of chemosensor. The receptor, a fluorophore and a spacer constitute the chemosensor [1]. From the porphyrinoid family, porphyrins and its analogues have been used as receptors extensively [2]. Recently, corroles have attracted scientific community attention because of the discovery of convenient protocols to synthesize corroles and their metal complexes [3]. Corroles shows unique properties and remarkable applications in catalysis, medicine, dye for solar energy conversion, optical and sensing applications due to direct pyrrole-pyrrole linkage [4,5]. The corroles are highly photostable in most of the solvents [6]. Corroles are highly fluorescent and show strong fluorescence emission with a large Stoke's shift and high radiative rate constants and absorption bands in the visible region [6]. As these chromophores are visible excitable so in vivo and environmental applications can be carried out in detail and results are quite positive. Corroles can be utilized as anion sensors due to their highly acidic nature of N—H proton [7]. It is evident from the literature that most of the research focused on the reactivity and properties of corroles. However, as compare to other porphyrinoid family, a very little development has been made for the use of corrole as anion sensors [5,8]. Because of diversification of fluoride anion in biological, chemical and environmental processes, the sensors for fluoride ion has become the domain of substantial interest for many years [9–10]. Its significant and important role in the prevention of dental caries, osteoporosis, psychiatric, hypnotic drugs and in nuclear technology to refine uranium is highly demanding [11–12]. On the other side high dose of fluoride into the living systems has resulted fluorosis, depression of thyroid activity and in disruption of immune system [13].

A large number of fluoride anion chemosensors are available in the literature due to emerging research interest in this field [14]. The molecule must be good hydrogen donor so that fluoride ion can polarize N—H bond which may result in deprotonation of an N—H bond [15–17]. This type of Bronston acid-base type mechanism will increase the fluoride concentration and will form highly stable complex $[HF_2]^-$ due to increase in basicity of fluoride [18]. But the interference from anions like benzoate, acetate and strongly basic anions such as sulphate and cyanide anion and tedious synthetic protocols are the serious drawbacks [19–20]. Mimicking of Boolean arithmetic with such type of cationic, anionic and photochemical switches are under progress to substitute silicon circuitry [21–22]. The painful fabrications and complexity to design molecule has brought success in several logic operations but advanced memory functions logic still is not so common [23–25]. It is clearly understood that anions has become the target of recent research due to their very significant and important participation into several bio-chemical processes. From our general chemistry knowledge, fluoride ion behaves like hard Lewis base, has smallest ionic radius and higher charge density. These unique properties of fluoride ions make it more interesting target for sensor design and its involvement in different technological, medical and biological processes. To overcome different technical challenges related to designing of fluoride ion sensors, herein, we present the synthesis, characterization, photophysical properties and sensing abilities of four novel free base A_2B corroles, 10-(pentaflurophenyl)-5, 15-bis(4-nitrophenyl)corrole (**1**), 10-(2, 6-Diflurophenyle)-5, 15-bis(4-nitrophenyl)corrole (**2**), 10-(3, 4, 5-triflurophenyle)-5, 15-bis(4-nitrophenyl)corrole (**3**), and

* Corresponding author.
E-mail address: anil_kumar@dce.ac.in (A. Kumar).

https://doi.org/10.1016/j.saa.2018.05.051
1386-1425/© 2018 Elsevier B.V. All rights reserved.

10-(4-methoxycarbonylphenyl)-5, 15-bis(4-nitrophenyl)corrole (**4**) (Scheme 1). The synthesized corrole, **1**–**4**, show good sensing ability selectively towards fluoride ion in the order **1** > **2** > **3** > **4**, while compound **1** also showed significant sensing ability for bromide ion.

2. Results and Discussions

In this work, synthesis of four novel free-base corroles, **1**–**4** (Scheme 1) were carried out. The standard protocols described in the literature were used for the synthesis of these corroles, **1**–**4** [26–27]. The synthesized corroles were characterized in detail using different spectroscopic techniques and their experimental procedures for, **1**–**4**, A_2B-corroles are systematized in the experimental section. The targeted A_2B corroles, contain either electron-withdrawing or electron donating groups on the *meso* phenyl ring and thus provide us an opportunity to study the effect of substituent on their photophysical properties and sensing abilities of corroles. Also, different synthetic parameters were optimized to obtain good yields for A_2B corroles, **1**–**4**, with the negligible amount of porphyrin.

The toluene solution of corroles, **1**–**4**, was used for the measurement of photophysical properties at ambient conditions and is tabulated in Table 1. The Q-type bands appeared around 590–655 nm and the Soret type bands were observed between 420 and 450 nm in toluene solution and are shown in Fig. 1. Usually, with the increasing value of dielectric constant of solvent, bathochromic shift in absorption spectra is observed but here in the case of bis(nitro)substituted corroles, the opposite trend was observed [4,28]. Ziegler et al. also noted the same behavior in the case of 10-methoxy-5,15-bis(nitrophenyl)corrole [4b]. So, this may be assumed in terms of uniqueness of moiety, A i.e. nitro, present on *meso* position of corroles which make the difference with absence of A moiety on free base corroles. The steady state fluorescence spectra of corroles, **1**–**4**, were recorded in toluene and bands appeared in the range of 673–690 nm. The recorded fluorescence emission spectra in toluene are shown in Fig. 2. The Stokes shift of 241 cm^{-1} was measured in the recorded emission intensity of all compounds, **1**–**4**.

This large Stokes shift may be due the change in the exited state nature as compared to ground state. The quantum yield for the corroles, **1**–**4**, are tabulated in Table 1 and were calculated by using the relation shown in Eq. (1) with respect to Rhodamine 6G ($\Phi = 0.95$).

$$\Phi_S = \Phi_R(I_S/I_R)\left(1-10^{-A}{}_R/1-10^{-A}{}_s\right)\left(n^2{}_s/n^2{}_R\right) \quad (1)$$

where Φ_R is the absolute quantum yield of standard reference Rhodamine 6G ($\Phi = 0.95$), I_S and I_R the integrated fluorescence intensity of the corrole and standard reference Rhodamine 6G (0.95), respectively. The absorbance for reference and sample are A_R and A_S, respectively. The term n_s and n_R represent the refractive index of solvent used. The emission intensity of corroles, **1**–**4** determined here are much higher than that of the earlier reported 5,15-bis(nitrophenyl) which indicates stronger contribution by nitrophenyl moiety via strong charge transfer [4b,28].

But opposite trend is observed in all these discussed herein novel corroles (**1**–**4**). In addition, the nanosecond range of lifetime measurements supports these trends. The highest and lowest lifetime were exhibited in the range of 4.90 (ns) to 3.61 (ns), respectively, and shown in Table 2. The observed decay curves were well fitted with multi-exponential model which is well explained by the Eq. (2) rather than mono-exponential function and shown in Fig. 3 [20].

$$I(t) = \alpha_1 \exp(-t/\tau_1) + \alpha_2 \exp(-t/\tau_2) \quad (2)$$

where τ_1 and τ_2 represent the lifetimes and α_1 and α_2 are the corresponding pre-exponential factors, respectively. The average lifetime, τ_{av} of corrole (**1**–**4**) is calculated using the Eq. (3) and are summarized in Table 2.

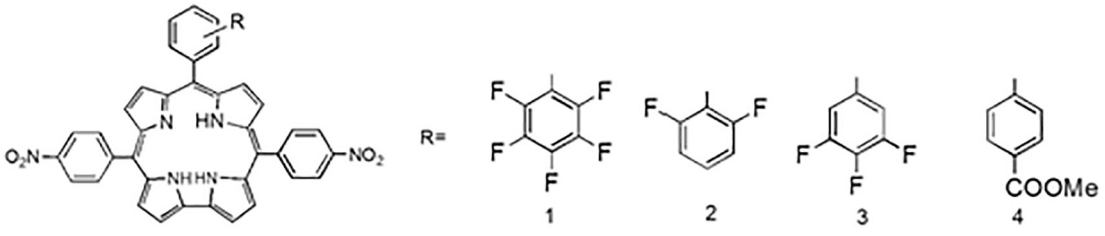
$$\tau_{av} = (\alpha_1\tau^2{}_1 + \alpha_2\tau^2{}_2)/(\alpha_1\tau_1/\alpha_2\tau_2) \quad (3)$$

It is well documented in the literature that polar solvent enhances the hydrogen bonding and charge transfer as well as stabilize the excited state which results in the decrease of emission life time and fluorescence [28]. The detection ability of ions at low concentration was also studied with the use of these compounds, **1**–**4**. The major problems like photostability, solubility, and spontaneous decomposition was taken into consideration for the fluorescence study [4b]. The very important factor, photostability of these corroles, **1**–**4**, in toluene solution for fluorescence quenching study was also considered. Fluorescence quenching investigation in toluene solution has been performed using corroles as a fluorophore [27]. The toluene as a solvent was chosen because of high solubility and no photodecomposition of corroles, **1**–**4**. To be very true, corrole solution in toluene remained stable for several months. Therefore, all the fluoride ion signaling experiments were carried out in toluene.

Corroles would be sensitive to basic anions because of highly acidic nature of inner N—H and fluoride ions would be the best candidate as compared to other basic ions. This may be due to the smaller size and higher electronegativity which in turn provide stronger interactions with the N—H of corroles through hydrogen bonding. This was clearly observed between fluoride ions and compound **1**–**4** [29]. The best performance in terms of sensing ability was observed for compound **1** as compared to **2**, **3** and **4**. These results can be justified by the presence of electron withdrawing groups on the *meso* phenyl group which can increase the acidic nature of inner N—H which in turn make these compounds highly sensitive to fluoride anions [30]. The spectral change in their ground and excited state shown in Fig. 4(a–d). The change in their absorption spectrum with addition of fluoride ion to compound **1**–**4** were also observed and are shown in Fig. 5a. With the addition of

Table 1
The calculated photophysical data of corroles (**1**–**4**) in toluene at 298 K.

Host	λ_{max} (ε) [nm ($\times 10^4$ cm^{-1} M^{-1})]	$\lambda^a{}_{em}$ (nm)	Stoke's shift	FWHM	QY%
1	439 (2.9), 593 (7.7), 649 (3.0).	673	241	39.60	0.76
2	439 (2.4), 598 (6.9), 655 (2.4)	680	244	43.78	0.80
3	442 (6.9), 589 (1.8), 648 (1.3).	687	249	49.01	0.77
4	425 (5.6), 445 (5.6), 590 (1.6), 645 (1.3).	689	244	47.13	0.51



Scheme 1. Compounds used for sensing ions.

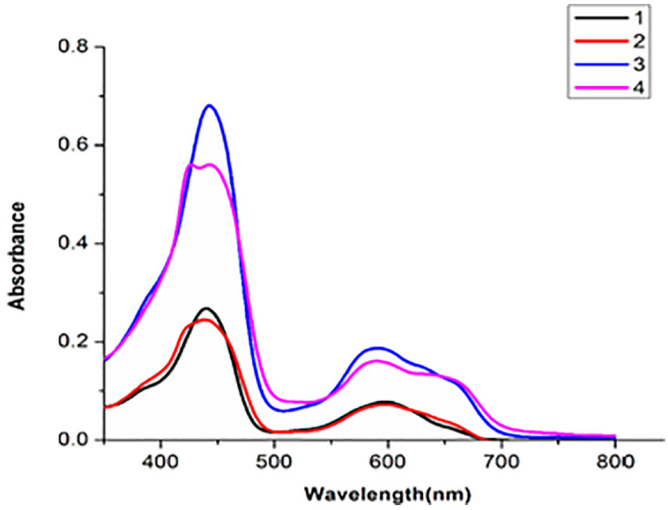


Fig. 1. Electronic spectra of toluene solution of synthesized free base corroles, **1**–**4**.

fluoride ion to compound **1**, an increase in the band at 666, 594 nm and decrease in the band at 498, 439 nm was observed. In addition to this change a red-shift from 594 nm to 666 nm was also collected ($\Delta_{\lambda max}$ - 72 nm) along with four isosbestic points at 628, 553, 460 and 422 nm. Similarly, in the case of **3**, an increase in band at 676 and 501 nm and decreases at 591, 441 nm (Fig. 5b) was also observed. Either hypsochromic or bathochromic shift was observed in the case of **3**. Surprisingly, compound **2** and **4** remains undetected by UV–visible spectroscopy.

A hypsochromic interfacial charge transfer (ICT) resulted with the addition of fluoride ion to host solutions which may be possibly due to interaction with the hydrogen atom of the inner N—H bond of the corrole. Also, it is well documented in the literature that fluoride ion quenches the fluorescence intensity through spin–orbital coupling effect [31].

A notable color change was observed in compounds **1**–**4** after the addition of fluoride ion from green to yellowish-brown. The molar extinction coefficient decrease upon binding with fluoride ion as well as hypsochromic shift of ~25 nm was measured in **1** and shown in Fig. 5a. The saturation point with the addition of excess of fluoride ion was observed and shown in the inset of Fig. 5a. It is clear from the Fig. 4 that fluorescence intensity get quenches upon addition of fluoride ion which may be understood in term of increase in non-radiative relaxation pathways from excited state to ground state. In other words, there

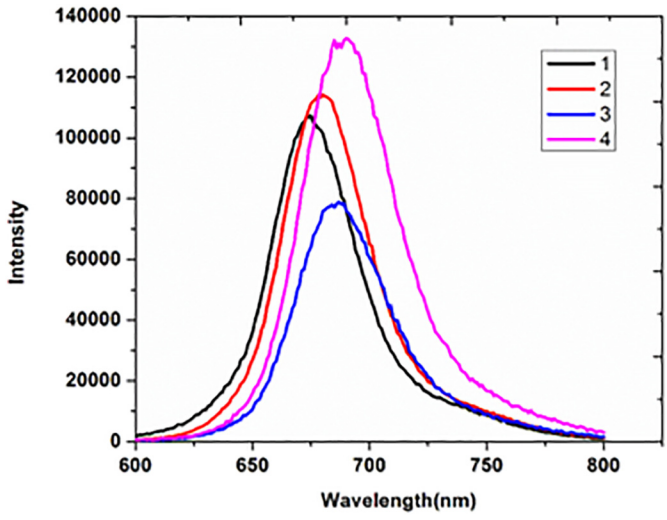


Fig. 2. Fluorescence emission spectra of toluene solution of **1**–**4** at 450 nm excitation.

Table 2
Lifetime and pre-exponential factors of corrole **1**–**4** at room temperature.

Host	τ_1 (ns)	τ_2 (ns)	B_1	B_2	τ_{Av} (ns)	χ^2
1	0.48 ± 0.049	4.47 ± 0.005	2.19	97.81	4.46	1.90
2	0.39 ± 0.03	3.79 ± 0.005	2.03	97.97	3.71	1.80
3	0.40 ± 0.036	3.88 ± 0.005	2.18	97.82	3.87	1.78
4	0.55 ± 0.05	3.69 ± 0.004	2.04	97.96	3.61	1.43

is increase in the interaction of fluoride ion with the corrole inner N—H. We could not see any new emission band thus definitely avoids the assumption of generation of a new fluorophore with the strong interaction of fluoride ion and corroles upon the addition of fluoride ion into the solution. Corrole **1** was found to exhibit the highest fluoride sensing ability among the corroles **1**–**4**. We observed a decreases in quenching efficacy in the order **1** > **2** > **3** > **4** under identical conditions and well presented in Fig. 4. The sensing ability of corrole **1**–**4** vary according to substituent B in all A_2B corroles because A is same in all the cases and only moiety B is different. The effect of substituent B is also reflected on its quantum yield and lifetime measurement.

In order to check the specific sensing ability of corroles **1**–**4**, a series of tetrabutylammonium (SO_4^{2-}, NO_3^-, HSO_4^-, CH_3COO^-, PO_4^{3-}), ions were used in the fluorescence titration experiment. The experiment results are well presented in the Fig. 6 which clearly shows that a quenching in the fluorescence intensity is lesser influenced by these anions relative to that of fluoride ion. The very peculiar selectivity and high affinity for fluoride ion was observed with the reported free base corroles **1**–**4**. From the Fig. 7, it is understood that significant quenching of the fluorescence intensity occurred with the addition of fluoride ion and at higher fluoride ion concentration the quenching is even more appreciable.

It is evident from the literature that Stern–Volmer plot is useful to study the nature of interaction for fluorescence quenching, so we thought of validating our assumption [31]. The Stern–Volmer plot in Fig. 7 shows the response (I_0/I-1) increases exponentially with increasing fluoride ion concentration. The obtained results indicate that chances of both k as well as static and dynamic interaction are operative in quenching. It is clear from the inset of Fig. 7 that only static interactions are responsible for 2.49×10^{-6} M (i.e. 1.0 equiv.). But further increase in the concentration of fluoride ion, the response increases exponentially. The Stern-Volmer quenching constant for fluoride ion interaction were determined for the compound **1**–**4**, $K_{sv} \times 10^6 \times L^{-1}$ and

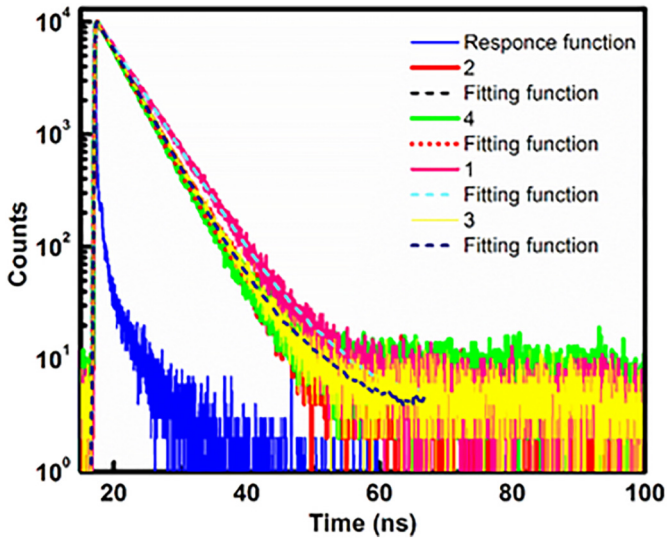


Fig. 3. The time-resolved photoluminescence decay curves (solid lines) together with simulated curves (dotted lines) of Corrole **1**–**4** measured at PL maxima. The excitation wavelength was 450 nm. The dotted line shows the response function.

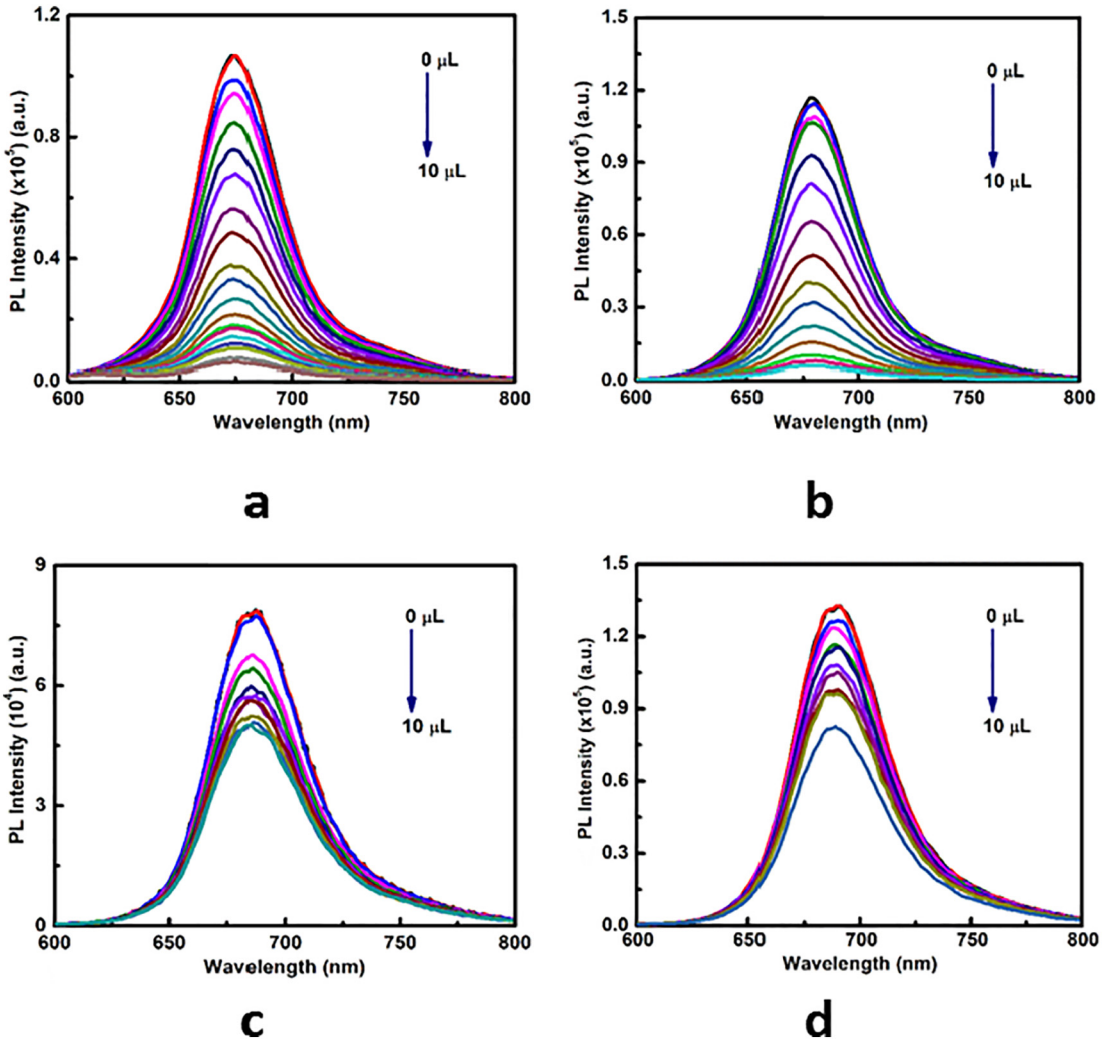


Fig. 4. Spectrofluorometric titration of **1**–**4** with the addition of fluoride ion in toluene of **1**, **2**, **3**, and **4**, and represented as **a**, **b**, **c** and **d**, respectively. Steady state fluorescence emission intensity changes were recorded for compound **1**, **2**, **3** and **4** at emission intensity 673, 680, 687 and 689 nm in air, respectively. Wavelength of excitation was 450 nm.

values are tabulated in Table 3. The obtained values show that interaction must take place inside the core via inner —NH. Further, the highest and lowest value of K_{sv} for compound **2** and **3** was calculated, respectively. As far as, the above discussed results, we concluded that **1** was the best sensor for fluoride ions but K_{sv} of compound **1** is lower than compound **2** and **4**. The detection limits, LOD, for the fluoride ion were also determined in the presence of compound **1**–**4** and shown in Table. 3. The absorption bands of compound **1**–**4** centered at 673, 678,

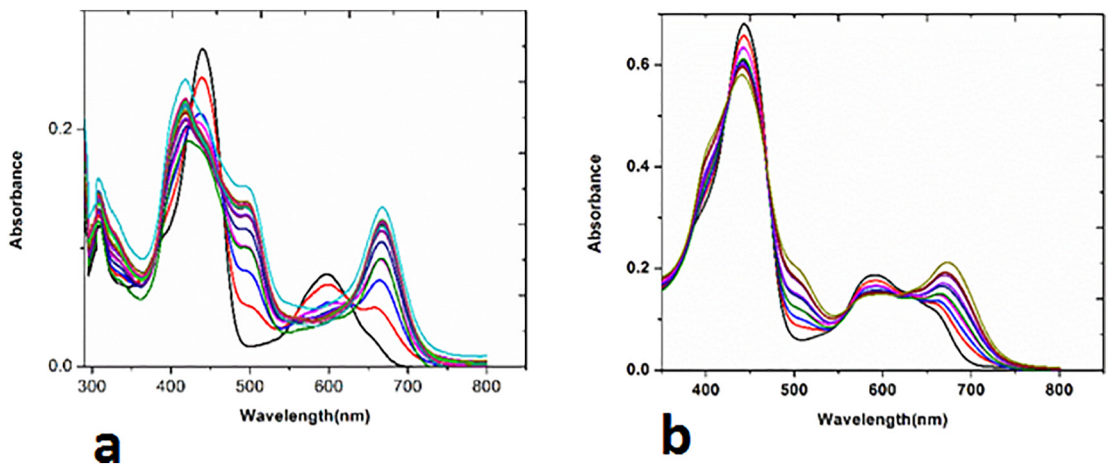


Fig. 5. The observed change in UV–visible absorption with the addition of different equivalents concertation of tetrabutylammonium fluoride into a dry toluene solution of (2.49×10^{-6} M) **1** and **3** and shown in a and b, respectively, at 298 K and in air. Compound **2** and **4** remained undetected by UV–visible spectrophotometer in this case.

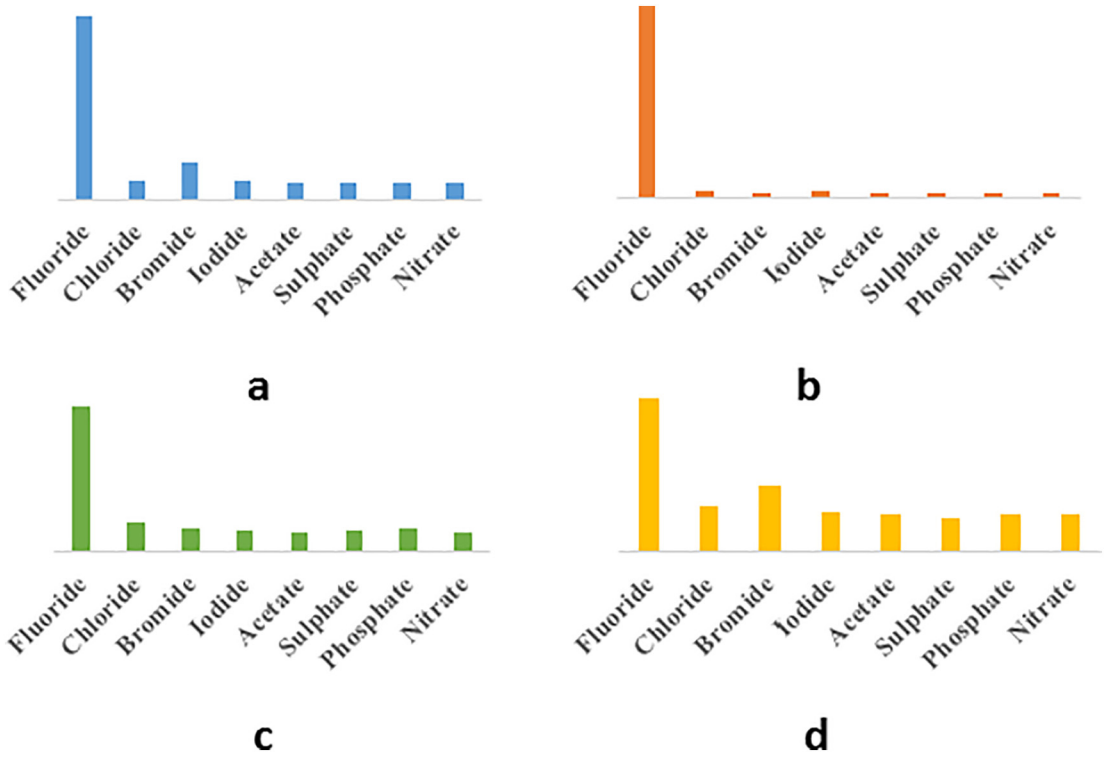


Fig. 6. The recorded fluorescence response of, **1**–**4** of 2.46×10^{-6} M in toluene represented as a–d, respectively.

687 and 690 nm the LOD calculated was 0.46, 0.44, 0.20, 0.30 respectively of, **1**–**4**, which is significantly higher than earlier report in the literature [27]. In all the free base corroles reported herein **1**–**4**, showed an excellent affinity for the fluoride ions with an unusual selectivity for fluoride ion. Further, it has been also discussed the interference of other analyte and results are well presented in Fig. 7 which clearly indicate that the response was found be best for fluoride ion than other environmentally relevant anions.

3. Conclusions

The newly synthesized nitrophenyl substituted freebase A_2B-corroles, **1**–**4** shows higher quantum yield than other porphyrinoid systems. These synthesized corroles have been demonstrated as promising fluorophores for the detection of fluoride ion. Quenching of the fluorescence intensity with the addition of fluoride ion was studied. These reported A_2B corroles, **1**–**4** shows higher quantum yield, emission intensity, lifetime and selective fluoride ion detection. As per our literature knowledge, we also conclude that, this is the third report of corroles to be used as fluoride ion sensor. These exciting results will motivate environmental scientific community to design the chemosensors for the detection of halide ions in the environment. Designing sensors for fluoride ion in aqueous media is the current foci of our laboratory.

4. Experimental and Physical Methods

All the synthesized corroles were purified over flash followed by column chromatography using silica gel as adsorbent. The time resolved fluorescence spectra and steady-state of the sample were coupled with lifetime attachment and data were collected in toluene solution by excitation at 450 nm using HORIBA spectrophotometer. Slits were set to 1 mm for the entrance, 1 mm for the intermediate and 1 mm for the exit. Quantum yield measurements experiments were performed in triplicate and calculation were made relative to Rhodamine 6G (Φ = 0.95). 450 nm laser with grated detector was used for lifetime measurements. All the binding experiments and anion detection titration were performed in toluene.

5. Synthesis

5.1. Synthesis of 5-(4-Nitrophenyl)dipyrromethane

5-(4-Nitrophenyl)dipyrromethane was synthesized from predistilled pyrrole and 4-nitrobenzaldehyde according to available reported method [18].

5.2. General Method for the Synthesis of Corroles

Corroles were synthesized as per the earlier reported methods in the literature [26–27]. 5-(4-Nitrophenyl)dipyrromethane (1 mmol) and respective aldehyde (0.5 mmol) were dissolved in MeOH (50 mL). Subsequently, a solution of HCl_{aq} (36%, 2.5 mL) in H_2O (50 mL) was added, and the reaction was stirred at room temperature for 1 h. The mixture was extracted with $CHCl_3$, and the organic layer was washed twice with H_2O, dried (Na_2SO_4), filtered, and diluted to 250 mL with $CHCl_3$. *p*-Chloranil (369 mg, 1.5 mmol) was added, and the mixture was stirred overnight at room temperature. Then the reaction mixture was passed through a small silica column to remove major portion of impurities. The filtrate was collected, evaporated to dryness and purified by column chromatography over silica gel. The analytical details for each compound are given below.

10-(Pentaflurophenyl)-5, 15-bis(4-nitrophenyl)corrole, **1**. The green color intense band was collected on separation with 4: 6; CH_2Cl_2: hexane as eluent. Yield = 38%. ^{1}HNMR ($CDCl_3$): δ: 9.075 (d, 2H, J = 3.6 Hz), 8.916 (d, 2H, J = 4.8 Hz), 8.725–8.704 (m, 8H), 8.614 (d, 2H, J = 4.4 Hz), 8.551–8.530 (m, 5H), 3.093 (s, 6H). MS (MALDI-TOF LD$^+$): m/z (%) observed: 706.16 $[M]^+$ (100); calculated: 706.14. UV–vis (toluene) λ_{max} (ε) [nm ($\times 10^4$ cm^{-1} M^{-1})]: 439 (2.9), 593 (7.7), 649 (3.0).

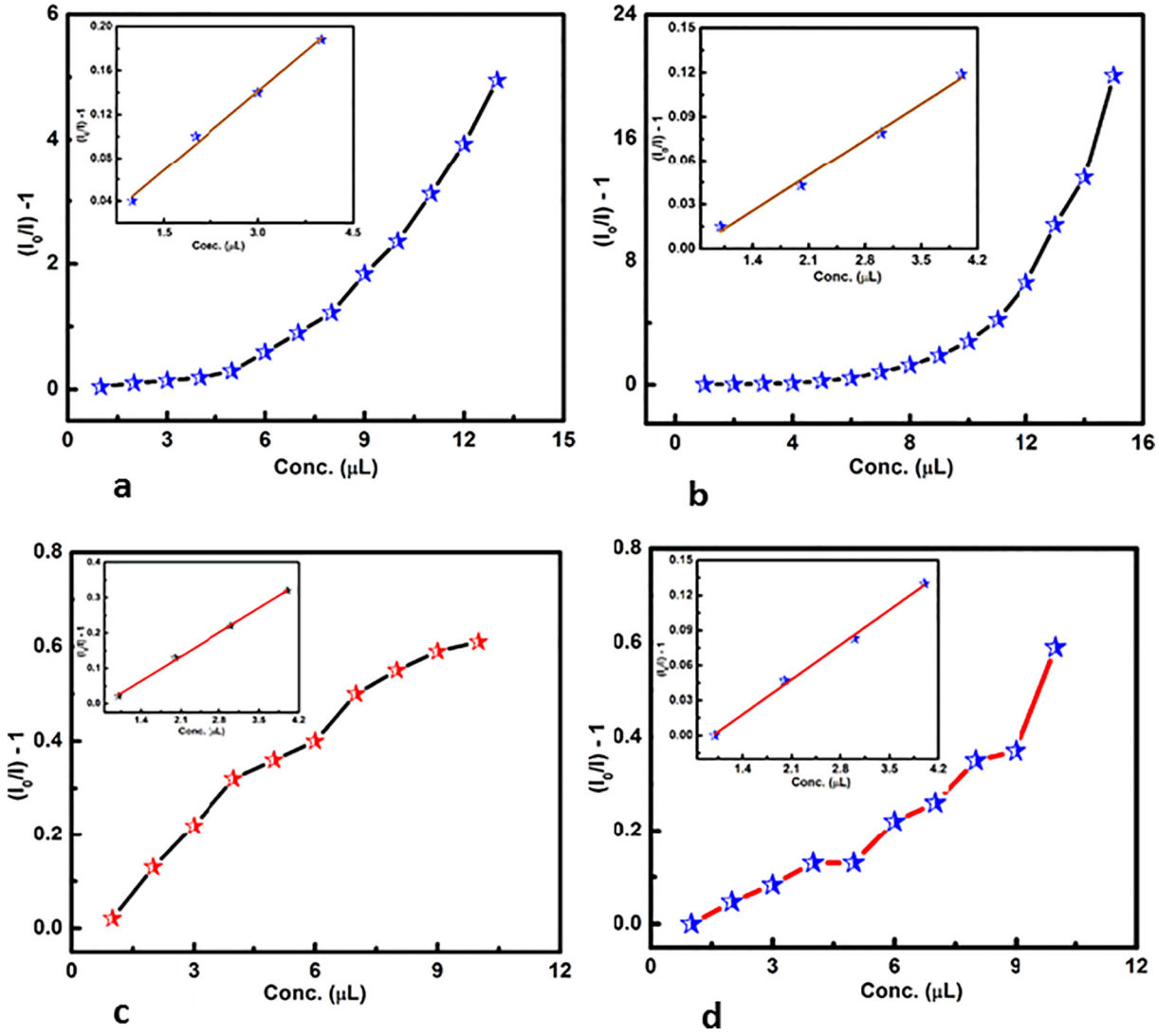


Fig. 7. Stern–Volmer plot of compound **1**–**4** showing positive deviation with increasing fluoride ion concentration. The figures a, b, c and d represent the corroles **1**, **2**, **3** and **4**, respectively. Conc. (μL) represents the amount of fluoride added in μL.

10-(2, 6-Diflurophenyle)-5, 15-bis(4-nitrophenyl)corrole, (**2**). The green color intense band was collected on separation with 9: 1; CH_2Cl_2: hexane as eluent. Yield = 35%. ^{1}H NMR ($CDCl_3$): δ: 9.05 (d, 2H, J = 4.4 Hz), 8.91 (s, 3H), 8.87 (m, 3H), 8.80 (d, 1H, J = 4.4 Hz), 8.71–864 (m, 5H), 8.59 (d, 4H, J = 4.4 Hz), 8.54 (d, 3H, J = 8.4 Hz), 8.41 (d, 1H, J = 8.4 Hz). MS (MALDI-TOF LD$^+$): m/z (%) observed: 652.13 $[M]^+$ (100); calculated: 652.17. UV–vis (toluene) λ_{max} (ε) [nm ($\times 10^4$ cm^{-1} M^{-1})]: 439 (2.4), 598 (6.9), 655 (2.4).

10-(3, 4, 5-Triflurophenyl)-5, 15-bis(4-nitrophenyl)corrole, (**3**). The green color intense band was collected on separation with 6:4; CH_2Cl_2: hexane as eluent. Yield = 29%. ^{1}H NMR ($CDCl_3$): δ: 9.097 (s, 2H), 8.902–8.882 (m, 3H), 8.819 (s, 1H), 8.719–8.630 (m, 7H), 8.548 (d, 3H, J = 10 Hz), 8.396 (d, 2H, J = 7.6 Hz). MS (MALDI-TOF LD$^+$): m/z (%) observed; 670.0 $[M]^+$: calculated: 670.0. UV–vis (toluene) λ_{max} (ε) [nm ($\times 10^4$ cm^{-1} M^{-1})]: 442 (6.9), 589 (1.8), 648 (1.3).

10-(4-Methoxycarbonylphenyl)-5, 15-bis(4-nitrophenyl)corrole, (**4**). The first intense green band was collected on separation with 3:1: CH_2Cl_2: hexane as eluent. Yield = 37%. ^{1}H NMR ($CDCl_3$): δ: 8.87 (d, 2H, J = 4.8 Hz), 8.70 (d, 2H, J = 8.4 Hz), 8.63–8.60 (m, 3H), 8.54 (d, 2H, J = 8.4 Hz), 8.45 (d, 2H, J = 8 Hz), 7.26 (d, 2H, J = 4 Hz), 4.11 (s, 3H). MS (MALDI-TOF LD$^+$): m/z (%): 674.12 $[M]^+$ (100): calculated: 674.19. UV–vis (toluene) λ_{max} (ε) [nm ($\times 10^4$ cm^{-1} M^{-1})]: 425 (5.6), 445 (5.6), 590 (1.6), 645 (1.3).

Table 3
Stern-Volmer constant and limit of detection.

Host	($K_{Sv} \times 10^6$ L^{-1})	LOD (10^6 L^{-1})	R^2
1	20.66	0.46	0.99
2	28.73	0.44	0.99
3	10.13	0.20	0.99
4	23.47	0.30	0.99

Conflict of Interest

There are no conflicts of interest to declare.

Acknowledgement

Omprakash Yadav and Atul Varshney thanks to CSIR and UGC for fellowship, respectively. Anil Kumar thanks to DTU for infrastructural facilities.

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.saa.2018.05.051.

References

[1] (a) J.P. Desvergne, A.W. Czarnik, NATO Science Series, Serie C: Mathematical and Physical Sciences, Kluwer Academic, London, 1997;
(b) C. Lodeiro, F. Pina, A.J. Parola, A. Bencini, A. Bianchi, C. Bazzicalupi, S. Ciattini, C. Giorgi, A. Masotti, B. Valtancoli, J.S. Melo, Inorg. Chem. 40 (2001) 6813–6819.
[2] Z. Xu, J. Yoon, D.R. Spring, Chem. Soc. Rev. 39 (2010) 1996–2006.
[3] (a) Z. Gross, N. Galili, I. Saltsman, Angew. Chem. Int. Ed. 38 (1999) 1427–1429;
(b) Z. Gross, N. Galili, L. Simkhovich, I. Saltsman, M. Botoshsnsky, D. Blaser, R. Boese, I. Goldberg, Org. Lett. 1 (1999) 599–602;
(c) R. Paolesse, L. Jaquinod, D.J. Nurco, S. Mini, F. Sagone, T. Boschi, K.M. Smith, Chem. Commun. (1999) 1307–1308;
(d) R. Paolesse, S. Nardis, F. Sagone, R.G. Khoury, J. Organomet. Chem. 66 (2001) 550–556.
[4] (a) B. Ventura, A. Degli Esposti, B. Koszarna, D.T. Gryko, L. Flamigni, New J. Chem. 29 (2005) 1559–1566;
(b) T. Ding, E.A. Alem, D.A. Mordarelli, C.J. Ziegler, J. Phys. Chem. A 109 (2005) 7411–7417;
(c) R. Paolesse, in: K.M. Kadish, K.M. Smith, R. Guilard (Eds.), The Porphyrin Handbook, vol. 2, Academic Press, New York 2000, pp. 202–203 , (ch. 11).
[5] L. Flamigni, D.T. Gryko, Chem. Soc. Rev. 38 (2009) 1635–1646.
[6] C.-L. He, F.-L. Ren, X.-B. Zhang, Z.-X. Han, Talanta 70 (2006) 364–369.
[7] I. Aviv-Harel, Z. Gross, Chem. Eur. J. 15 (2009) 8382–8394.
[8] D.T. Gryko, J.P. Fox, D.P. Goldberg, J. Porphyrins Phthalocyanines 8 (2004) 1091–1105.
[9] (a) Y. Zhou, J.F. Zhang, J. Yoon, Chem. Rev. 114 (2014) 5511–5571;
(b) E. Galibraith, T.D. James, Chem. Soc. Rev. 39 (2010) 3831–3842.
[10] (a) P.A. Gale, Chem. Soc. Rev. 39 (2010) 3746–3771;
(b) J.W. Jones, L.N. Zakharov, A.L. Rheingold, H.W. Gibson, J. Am. Chem. Soc. 124 (2002) 13378–13379;
(c) F.M. Pfeffer, T. Gunnlaugsson, P. Jensen, P.E. Kruger, Org. Lett. 7 (2005) 5357–5360.
[11] (a) S. Guha, S. Saha, J. Am. Chem. Soc. 132 (2010) 17674–17677;
(b) H. Miyaji, W. Sato, J.L. Sessler, Angew. Chem. Int. Ed. 39 (2000) 1777–1780.
[12] (a) G. Waldbott, Clin. Toxicol. 18 (1981) 531–541;
(b) B.L. Riggs, Bone Miner. Res. (1984) 366–393;
(c) S. Matuso, K.-i. Kiyomiya, M. Kurebe, Arch. Toxicol. 72 (1998) 798–806.
[13] (a) R.H. Dreisbuch, Handbook of Poisoning, Lange Medical, Los Altos, CA, 1980 578;
(b) A. Wiseman, Handbook of Experimental Pharmacology XX/2, Springer, Berlin, 1970 48 (Part 2).
[14] (a) J.M. Linares, D. Powell, J.K. Bowman, Coord. Chem. Rev. 240 (2003) 57–75;
(b) P.D. Beer, P.A. Gale, Angew. Chem. Int. Ed. 40 (2001) 486–516.
[15] (a) P.A. Gale, N. Busschaert, C.J.E. Haynes, L.E. Karagiannidis, I.L. Kirby, Chem. Soc. Rev. 43 (2014) 205–241;
(b) M.A. Hossain, S.A. Kang, J.A. Kut, V.W. Day, K. Bowman-James, Inorg. Chem. 51 (2012) 4833–4840.
[16] (a) C. Perez-Casas, A.K. Yatsimirsky, J. Organomet. Chem. 73 (2008) 2275–2284;
(b) F.Y. Wu, Z.C. Wen, N. Zhou, Y.F. Zhao, Y.B. Jiang, Org. Lett. 4 (2002) 3203–3205.
[17] (a) Q.-Y. Cao, M. Li, L. Zhou, Z.-W. Wang, RSC Adv. 4 (2014) 4041–4046;
(b) X.J. Peng, Y.K. Wu, J.L. Fan, M.Z. Tian, K.L. Han, J. Organomet. Chem. 70 (2005) 10524–10531.
[18] (a) Q. Wang, Y. Xie, Y. Ding, X. Lie, W. Zhu, Chem. Commun. 46 (2010) 3669–3671;
(b) K. Choi, A.D. Hamilton, Angew. Chem. Int. Ed. 40 (2001) 3912–3915;
(c) V. Amendola, L. Fabbrizzi, L. Mosca, Chem. Soc. Rev. 39 (2010) 3889–3915;
(d) N. Ahmed, V. Suresh, B. Shirinfar, I. Geronimo, A. Bist, I.C. Hwang, K.S. Kim, Org. Biomol. Chem. 10 (2012) 2094–2100.
[19] (a) H.S. Jung, H.J. Kim, J. Vicens, J.S. Kim, Tetrahedron Lett. 50 (2009) 983–987;
(b) M. Cametti, K. Rissanen, Chem. Commun. (2009) 2809–2829.
[20] (a) A.P. de Silva, M.R. James, B.O.F. McKinney, D.A. Pears, S.M. Weir, Nat. Mater. 5 (2006) 787–790;
(b) A.P. de Silva, H.Q.N. Gunaratne, C.P. McCoy, Nature 364 (1993) 42–44;
(c) J.-M. Lehn, Supramolecular Chemistry: Concepts and Perspectives, Wiley-VCH, Weinheim, New York, 1995.
[21] (a) S. Uchiyama, N. Kawai, A.P. de Silva, K. Iwai, J. Am. Chem. Soc. 126 (2004) 3032–3033;
(b) J. Andreasson, U. Pischel, Chem. Soc. Rev. 39 (2010) 174–188;
(c) W. Li, Y. Yang, H. Yan, Y. Liu, Nano Lett. 13 (2013) 2980–2988.
[22] (a) V. Balzani, A. Credi, M. Venturi, Molecular Devices and Machines - A Journey into the Nano World, Wiley-VCH, Weinheim, 2003 235;
(b) S. Sreejith, K.P. Divya, T.K. Manojkumar, A. Ajayaghosh, Chem. Asian. J. 6 (2011) 430–437;
(c) A.P. de Silva, N.D. Mcclenaghan, Chem. Eur. J. 18 (2012) 3549–3558.
[23] (a) W. Lu, M. Zhang, K. Liu, B. Fan, Z. Xia, L. Jiang, Sensors Actuators B 160 (2011) 1005–1010;
(b) R. Martins, P. Barquinha, L. Pereira, N. Correia, G. Goncalves, I. Ferreira, E. Fortunato, Appl. Phys. Lett. 203501 (2008) 93;
(c) F. Galindo, J.C. Lima, S.V. Luis, A.J. Parola, F. Pina, Adv. Funct. Mater. 15 (2005) 541–545.
[24] (a) G. Zong, G. Lu, J. Phys. Chem. C 113 (2009) 2541–2546;
(b) Y.I. Liu, M. Li, Q. Zhao, H. Wu, K. Huang, F. Li, Inorg. Chem. 50 (2011) 5969–5977;
(c) B. Kumar, M.A. Kaloo, A.R. Sekhar, J. Sankar, Dalton Trans. 43 (2014) 16164–16168.
[25] D.T. Gryko, K. Jadach, J. Organomet. Chem. 66 (2001) 4267–4275.
[26] (a) R. Paolesse, L. Jaquinod, D.J. Nurco, S. Mini, F. Sagone, T. Boschi, K.M. Smith, Chem. Commun. (14) (1999) 1307–1308;
(b) R. Paolesse, S. Nardis, F. Sagone, R.G. Khoury, J. Organomet. Chem. 66 (2001) 550–556;
(c) Z. Gross, N. Galili, L. Simkhovich, I. Saltsman, M. Botoshansky, D. Blaser, R. Boese, I. Goldberg, J. Organomet. Chem. (4) (1999) 599–602;
(d) Z. Gross, N. Galili, I. Saltsman, Angew. Chem. Int. Ed. 38 (1999) 1427–1429.
[27] (a) A. Pariyar, S. Bose, S. Singh Chhetri, A. Narayan, B. Biswas, P. Bandyopadhyay, Dalton Trans. 41 (2012) 3826–3838;
(b) C.I.M. Santos, E. Oliveira, B.J.F.B. Barata, M.M.F. Faustino, J.A.S. Cavaleiro, M. Grac, P.M.S. Neves, C. Lodeiro, J. Mater. Chem. 22 (2012) 1381;
(c) C.I.M. Santos, E. Oliveira, J.F. Lodeiro, J.F.B. Barata, S.M. Santos, M.A.F. Faustino, J.A.S. Cavaleiro, M. Graça, P.M.S. Neves, Carlos Lodeiro, Inorg. Chem. 52 (2013) 8564–8572.
[28] A. Mahammed, J.J. Weaver, H.B. Gray, M. Abdelas, Z. Gross, Tetrahedron Lett. 44 (2003) 2077–2079.
[29] B. Valeur, Molecular Fluorescence: Principles and Applications, Wiley-VCH Verlag GmbH, 2001.
[30] (a) D.S. McClure, J. Chem. Phys. 20 (1952) 682–686;
(b) P. Kumar, V. Kumar, R. Gupta, RSC Adv. 5 (2015) 97874–97882.
[31] (a). R.K. Ratnesh, M.S. Mehata, Spectrochim. Acta A Mol. Biomol. Spectrosc. 179 (2017) 201–210;
(b) R.K. Ratnesh, M.S. Mehata, AIP Adv. 15 (2015) 097–114;
(c) M.S. Mehata, M. Majumder, B. Mallik, N. Ohta, J. Phys. Chem. C 114 (2010) 15594–15601;
(d) M.S. Mehata, Sci. Rep. 5 (2015) (12056-11);
(e) M.S. Mehata, Appl. Phys. Lett. 100 (2012) (151908-4);
(f) R.K. Ratnesh, M.S. Mehata, Opt. Mater. 64 (2017) 250–256.

National Conference on Recent Advances in Computer Science and IT (NCRACIT)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 1 | ISSN: 2456-3307

A Comprehensive Study of Text Summarization Algorithms

Yash Dhankhar[1], Indu Bala[2], Swati Singh[3], Sunil Dalal[4]

[1]Computer Science and Engineering, BabaMastnath Engineering College, Rohtak, Haryana, India
[2]Information Systems, Delhi Technological University, Delhi, India
[3]Computer Science and Engineering, Delhi Technological University, Delhi, India
[4]Assistant Professor, BGSB University, Rajouri, J&K,India
yashdhankhar92@gmail.com[1], induyadav330@gmail.com[2], swatisngh337@gmail.com[3], sunildalal57@gmail.com[4]

ABSTRACT

This document provides some minimal guidelines (and requirements) for writing a research paper. Issues related to the contents, originality, contributions, organization, bibliographic information, and writing style are briefly covered. Evaluation criteria and due dates for the research paper are also provided.

Keywords : Research Paper, Technical Writing, Science and Technology

I. INTRODUCTION

Text summarization is a method to produce a concise and important piece of information from a larger set of text which can be a text document, an article or a blog. Text Summarization aims to provide a summary of given text while preserving its information and intent. The summary is a small piece of information that describes a set of paragraphs or documents. Summary generated is generally less than forty percent of the original text data and it should be even less than that in the case of large datasets. The summary should retain the important data present in the document, should be controllable, short and succinct. Summarization of text data is done in many ways depending upon the various parameters based on the position and format of words and sentences. Automatic Text Summarization [1] accumulates the data from several documents to present the final shorter piece of information as a result, which is shorter, informative and preserves the real intent of information. These small summarized versions save valuable time by presenting unambiguous important information. With the increasing amount of digital data, it has become difficult to retrieve the needed and concise information. Automatic text summarization caters to the very need of the time.

There are methods, which are helpful to produce a summary. First Division, which categorizes the summarization approaches, is based on the content of the summary produced. There are two approaches- Extraction and Abstraction [2]. As the name suggests, Extraction is domain independent, it mainly aims at finding out the important sentences and later presenting a set of important sentences as Summary. On the contrary, Abstraction is domain dependent, it processes the available information and new sentences are prepared by understanding the content, also considers human knowledge by preparing the goal to produce a summary.

Applications of Text Summarization: Automatic text summarization has many important applications. One of the important application is in Medical area, where a lot of unclassified information is available and many times a medical associate is required to find about some information specific to a medical condition for research or diagnostic purpose from the

large heap of documents. Finding out relevant information involves the reading of numerous documents and problem/patient's records. Summarization specifically personalized to the medical area is very useful, as it not only saves time but increases the efficiency of a medical expert.In legal processes, a typical case study involves consideration of loads of information consisting of law expertise books and numerous related previous judicial case studies, thus leading to an overload of information. The legal experts perform a tedious and responsible task, their time and resources are expensive. To find out an important piece of information unambiguously and in less time is desirable to cater the needs of fast and correct court decisions. For Research Purposes, hundreds of research papers need to be considered in any research domain to find out a specification. In this way to know what lies inside of the paper, researchers need to read more than the abstract but less than the paper, so summarization may be applied to get the customized summary of the content by applying the desirable method. On the internet, Summarization is used in multiple applications. Various newspaper sites and related apps provide everyday news with the use of summarization in order to save time and space while keeping the important key information. Mainly editorials are summarized while keeping the intent of author intact. Further, there are also applications for mobile devices like smart phones, tablets etc. they include small screen area and time available to read. In the corporate world, 'meeting minutes' need to be read in small time and associated documents need to be looked into before next meeting without the help of human and other resources. For blind people, a lot of time can be saved by readers while reading to them by giving them an important piece of information instead of the whole document. The paper is structured as follow: Section 2 gives an insight into the related work done in text summarization. Section 3 gives an overview of automatic text summarization. Section 4 concludes the paper and gives the future scope in this research area.

II. RELATED WORK

A lot of good work has been done on the text Summarization. In most of the Summarizers, Sentences are considered as a feature vector[] and various algorithms are applied depending upon the position of Sentences, Vocabulary intersections, title distribution and the type of data. Apart from the sentence related data other features include the structure of the document and popularity of the content. Most of the work has been done in the extraction domain, but various different ideas have been explored like multiple document summaries, language based summarizers etc. In 1955, Henry Peter Luhn, IBM inventor first published a paper entitled 'A new method of recording and searching information' (Luhn, 1953) [3]. He developed many Information retrieval applications. Later in 1969, Edmundson described a new extraction method based on extraction using three components: pragmatic words(cue words), structural indicators and topic heading words[4]. In 1980's AI methods came into existence for summarization using hybrid approaches for different types of summarization i.e. multiple documents, multimedia etc. One of the applications is KWIC (Keyword in Context) by using three fundamental elements: Keyword, title, and context. In the last two decades various new and hybrid methods have been described. TextRank[5], cluster based[6], Rhetoric based[7], Topic models[8], ILP based method[11] etc.

In an Abstractive domain, as new sentences need to be designed, it needs a deeper analysis of the original text information. It involves an understanding of the text by linguistic methods[9] to provide an interpretation to match the level of human generated summary. There are two main approaches for doing this i.e. structure based and semantic based approach. In Structure based approach, most weighted data is encoded by cognitive schemas[10]. Structures such as a tree, ontology, lead and body phrase structures are the schemas mostly used for structured approach. In the

second approach, semantic based uses a Natural Language generation system to process the semantic information to categorize the grammar variants such as noun and verbs by processing linguistic data. To achieve a true abstractive summarization is still a dream.

III. OVERVIEW OF AUTOMATIC TEXT SUMMARIZATION

A. Features

In order to decide the degree of importance a sentence to include or exclude from the final summary formation, a list of features used by the researchers are listed below:

1) Term Frequency [19]: Frequency of a word is measured for the whole source document. Then, the scores are assigned to each sentence based upon the number of frequent words belonging to the particular sentence. Sentences with highest weights are considered for final summary. TF IDF is widely used for calculation of word frequency.
2) Location: In a text document, the position of a sentence also tells about its relevance in the summary. While calculating the weighted Score of a sentence certain sentences are weighted higher than others.
3) Cue Method: Sentences including words that adds to limitations or advantages of the content are weighted high i.e. “in summary”, “significantly”, “describes” “concludes” are cue words.
4) Title/ Headline[20]: words included in the topic or theme of the content are considered relevant. Sentences which includes these words are assumed to be important for the summary. Some constant weight is added while weight calculation.
5) Sentence Length[19]: Number of words in a sentence defines the length of sentence, which is a factor in deciding sentence relevance in the final summary.
6) Proper noun [19]: Source document sentences, which include proper nouns, are assumed suitable for the final summary.
7) Proximity [19]: To identify relations among words or entities, distance between them is considered an important factor.
8) Similarity: To find out the relevance of a sentence in a document we calculate similarity among sentence and other sentences of the source document.

B. Summarization Approaches

The summarization can be performed on single documents and multiple documents as well.

- Single Document Summarization: In single document summarization[16], one source document is analyzed and processed to generate a quality summary. Both the approaches, extractive as well as abstractive can be applied on a single text document.
- Multi Document Summarization: Multi document summarization is a technique which involves the information extraction from more than one document. Multiple source documents are analyzed and evaluated to generate an important and non-redundant piece of information. Multiple document summarization [17] technique came after single document summarization to cater the needs when we need to concise data, which is distributed in multiple files. News on the internet is based on the web based clustering systems.

C. Methods

Two methods, extractive and abstractive are discussed here.

1) Extractive Summaries

Extractive summaries are simple to form as they only include few important sentences from the text document. They decide the importance of sentences in a document and decides to include the most informative sentences, paragraphs etc. in the final summarized result.

2) Abstractive Summaries

Abstractive summaries are prepared with a combination of newly formed sentences by analyzing a set of important information. New formed sentence should be coherent and complete. Abstractive summaries are generated by proper understanding the source document and then forming new sentences. It produces a representation of internal semantic details, then uses the natural language techniques for the final summary generation.

D. Types

Query based summaries and generic summaries are the two types of summarization types.

1) Query Based Summaries

In Query based summaries [18]; the final summary is generated based on query raised by a user. This technique can be applied on the single document as well as multiple documents. The relevance of a sentence for the final summary result is calculated based upon the frequency of words in a document. A sentence in the original document, which includes the keywords provided in a query by the user, is scored high than others. Sentences with the high scores are suitable for final summary.

2) Generic Summaries

Generic summaries provide a complete review of the source document unlike query based technique only caters to the query of the user. For the content overview, generic summaries are suitable. This aims to identify the key topics and decrease the redundancy to a possible minimum. Generic summaries categorize and describe the main idea of the source content.

E. Summary Techniques

Four well-known techniques are used in summarization.

- Semantic and Syntactic (Rule-based)
- Statistical Technique
- Clustering Technique
- Machine Learning Technique

1) Semantic and Syntactic Analysis

Rule-based technique is used to find and present the association among different sentences by applying on source content for text summarization. These can be categorized as following:

- Graph Representation
- Lexical Chains
- NLP (Natural Language Processing)

- The graph representation is done during summarization by lexical graphs, sentences are represented as Weighted graphs, unweighted graphs, graph matching etc are tasks performed during summarization process.
- Lexical chains are used for building chains of identified units for summarization with the help of co-reference chains and lexical semantics etc.
- Natural Language information processes language data to extract information also uses part of speech for summary production. There are two techniques for summarization under Natural Language Processing listed below:
  1) Plain text Summarization
  2) Multilingual Summarization

In Plain text summaries, resultant summary is in the same natural language but in multilingual text summarization [21] resultant summary may be in one of the languages in which sources are written or may be a totally different language. It came into existence in 2005. Evans (2005) described the need of summarization in a particular language from different sources available in different languages.

This technique is still in research stage but it has features that are very useful in news reporting where data is combined from different foreign news agencies and summarized in language used in a region.

2) Statistical Technique

For extraction of relevant information, some systems use Statistical Techniques. This technique uses statistical methods generally applied with Binomial Distribution, sentence compression and calculated scores. This technique is used by Hidden Markov model.

Conroy and O'Leary [24] employed statistical technique by hidden Markov model approach for summarization of plain text documents. A sequential Model was prepared for the evaluation of local independence.

This system has three key parameters as the length of the sentence in processing, the position of the sentence in document and likeliness of key terms in the sentence being evaluated.

3) Clustering Technique

When multiple objects are grouped together based upon their properties and characteristics, this process is termed as Clustering[6]. A cluster consists of the objects having similar properties. In text summarization, we use clustering to group similar type of sentences together. In a document different topics are arranged in a specific ordering. In this technique, firstly clusters are generated and then sentences are selected. The sentence is also chosen based upon the location or position of the sentence in a document. A score of a sentence increases if it has multiple occurrence hence higher probability of selection in the final summary.

4) Machine Learning Technique

Automatic text summarization can be effectively done by Machine learning techniques. Some of the machine learning approaches are discussed as follows.

a) Naive Bayes Approach : Kupiec (25) described a method for summarization in which a classification function known as naïve Bayes classifier is used which is responsible for the each sentence to be a part of the summary.
b) Rich features and Decision Trees: Lin and Hovy (27) used "sentence position" in which a weight is provided to sentence based upon its position in the text. This method is also known as position method.
c) Log Linear Models : Osbrone (2002) described the Log Linear model approach[28] for the plain text summarization. This approach is different than the previous approaches which always assumed feature independence. The system showed that this approach is better than naïve Bayes classifier approach.
d) Neural Networks : Svore (29) produced an algorithm based upon neural networks and used the third party features like dataset to resolve the problem of extractive summarization.

F. Comparative Study

As Text summarization first approach came in the 1950s since then many new approaches and techniques have been implemented and exercised. Different techniques involve a specific set of feature selection and the content on which the algorithm is applied, We have done a comparative analysis of few techniques in the table listed below:

Author/Year	Method	Features/ Content Selection	Technique Used	Summarization Approach
1995 Julian Kupiec[25]	algebraic method	like length, the position of words, uppercase words	using a naïve-bayes classifier	Extractive Summarization
1997 ChinYew Lin[41]	algebraic method	the position of sentences	Rich Features and decision trees	Extractive Summarization
1999 Eduard Hovy[31]	symbolic word knowledge	concepts relevancy	NLP processing	Single Document Summarization(A)
2005 S.P Yong[42]	Text pre-processing and	Keywords Extraction	used neural networks	Abstractive Summariz

	subsyste m	Summary productio n		ation
1984 Ruqaiya Hasan[43]	Coheren ce relation	similarity chains	lexical cohesion	Single Documen t Summariz ation(A)
1988 William C.Mann [44]	Tree based	to encode the terminal nodes of a tree	RST (rhetorical structure theory)	Abstracti ve Summariz ation
1997 Branimir Boguraev[45]	Saliency based content characte rization	rank the important sentences	Ranking algorithm	Extractiv e Summariz ation
2010 Li Chengche ng [46]	rhetoric relations	candidate sentence	RST (rhetorical structure theory)	Abstracti ve Summariz ation
Xiaojun Wan in 2008 [47] used graph based method by introduci ng	used graph based method	The two-link graph for both sentences and document s	Graph based method	Multiple Documen t Summariz ation(A)
2012 Tiedan Zhu [48]	Sentence closeness Paramet er	Logical closeness to document	Sentence Co-relation Method	Multiple Documen t Summariz ation€

IV. CONCLUSION

Automatic Text Summarization is used to get an important piece of text from a larger document. A large number of algorithms designed and implemented to get a good, coherent and non-redundant summary a little similar to the human prepared summary. Simple single document extractive algorithms have given better results in different domains as compared to abstractive summarization algorithms. Extractive summarizers are used to select the important set of sentences from the source document based on top scoring Sentence-ranking method. Although, By performing Automatic Text Summarization to get a gist of the input text documents equivalent to human interpreted summary is not yet fulfilled, but by improving the existing algorithms, the value of evaluation metrics is increasing. With the rapid increase in the electronic data on internet and less time to read the documents based on a similar topic has called a need to design accurate and efficient Multi- document summarization systems. As research on text summarization started 50 years ago and a lot of work has been done in the extractive area in both the single and multiple document domains but there is still a long path to cover in this field. Abstractive summarizers aim to import more information in a single sentence rather than include the sentence as a whole. Multi-document Abstractive Summarization is the area which is needed to be explored.

Over time, attention has drifted from summarizing scientific articles to news articles, electronic mail messages, advertisements, and blogs. Domain associated summarizes can be a solution to get more accurate summaries. Medical and Legal matters domain can be highly benefitted from this area of research even if they focus only on small details related to a general summarization process and not on building an entire domain dependent summarization system.

II. REFERENCES

[1] D.K. Gaikwad and C.N. Mahender "A Review Paper on Text Summarization", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016,154-160

[2] Radev, D. R., Hovy, E., and McKeown, K. (2002) "Introduction to the special issue on summarization." Computational Linguistics., 28(4):399-408

[3] Luhn, H. P. (1958) " The automatic creation of literature abstracts". IBM Journal of Research Development, 2(2):159–165

[4] Edmundson, H. P. (1969) " New methods in automatic extracting". Journal of the ACM, 16(2):264–285.

[5] R.Mihalcea, and P.Tarau, "TextRank: Bringing Order into Texts." In Proceedingsof Empirical Methods in Natural Language Processing (EMNLP). pp. 404-411. 2004.

[6] Z.Pei-ying, and L.Cun-he, "Automatic Text Summarization based on Sentences Clustering and Extraction," Proceeding of the 2nd IEEE International Conference on Computer Science and Information Technology. pp. 167-170. 2009

[7] 20IOy International Conference on Computer Application and System Modeling (ICCASM 2010) Automatic Text Summarization Based On Rhetorical Structure Theory Li Chengcheng 595-598[8] D. Blei, A. Ng, and M. Jordan " Latent Dirichlet allocation". In Journal of Machine Learning Research, 3:993–1022, January2003.

[8] Barzilay, R. and Elhadad, M. (1997). "Using lexical chains for text summarization." in Proceedings ISTS'97. pg. 38-41

[9] Radev, D. R. and McKeown, K. (1998) "Generating natural language summaries from multiple on-line sources." Computational Linguistics, 24(3):469–500

[10] S. Banerjee, P.Mitra and K. Sugiyama " Multi-Document Abstractive Summarization Using ILP Based Multi-Sentence Compression"in Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015)

[11] Lin, C.-Y. (2004). "Rouge: A package for automatic evaluation of summaries." In Proceedings of the ACL-04 Workshop, pages 74–81, Barcelona, Spain

[12] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu "BLEU: a Method for Automatic Evaluation of Machine Translation" in Computational Linguistics (ACL), Philadelphia, July 2002, pp. 311-318

[13] S. Brin and L. Page "The PageRank Citation Ranking:Bringing Order to the Web" in 1999

[14] Mc Keown, K. R. and Radev, D. R. (1995). "Generating summaries of multiple news articles." in Proceedings of SIGIR '95, pages 74–82, Seattle, Washington.

[15] Jagadeesh J, Prasad Pingali, Vasudeva Varma "Sentence Extraction Based Single Document Summarization" Workshop on Document Summarization, 19th and 20th March, 2005, IIIT Allahabad

[16] Kamal Sarkar, "Sentence Clustering-based Summarization of Multiple Text Documents", TECHNIA – International Journal of Computing Science and Communication Technologies, vol. 2, no. 1, Jul. 2009.

[17] F. Canan Pembe and Tunga Güngör, "Automated Query-biased and Structure-preserving Text Summarization on Web Documents," in Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications, İstanbul, June 2007.

[18] Reeve Lawrence H., Han Hyoil, Nagori Saya V., Yang Jonathan C., Schwimmer Tamara A., Brooks Ari D., "Concept Frequency Distribution in Biomedical Text Summarization", ACM 15th Conference on Information and Knowledge Management (CIKM), Arlington, VA, USA,2006.

[19] Khan Atif, Salim Naomie, "A review on abstractive summarization Methods", Journal of Theoretical and Applied Information Technology, 2014, Vol. 59

[20] Evans, D. K. (2005). "Similarity-based multilingual multi-document summarization." Technical Report CUCS-014-05, Columbia University.

[21] Edmundson, H. P. (1969). "New methods in automatic extracting." Journal of the ACM, 16(2):264–285.

[22] Martins, Camilla Brandel and Lucia Helena Machado Rino. "Revisiting UNLSumm: Improvement Through a Case Study." (2002).

[23] Conroy, J. M. and O'leary, D. P. (2001). "Text summarization via hidden markov models." In Proceedings of SIGIR '01, pages 406–407, New York, NY, USA

[24] Kupiec, J., Pedersen, J., and Chen, F. (1995). "A trainable document summarizer." In Proceedings SIGIR '95, pages 68–73, New York, NY, USA.

[25] Aone, C., Okurowski, M. E., Gorlinsky, J., and Larsen, B. (1999). "A trainable summarizer with knowledge acquired from robust nlp techniques".pages 71–80

[26] Lin, C.-Y. and Hovy, E. (1997). "Identifying topics by position." In Proceedings of the Fifth conference on Applied natural language processing, pages 283–290, San Francisco, CA, USA.

[27] Osborne, M. (2002). Using maximum entropy for sentence extraction. In Proceedings of the ACL'02 Workshop on Automatic Summarization, pages 1–8, Morristown, NJ, USA

[28] Svore, K., Vanderwende, L., and Burges, C. (2007). "Enhancing single-document summarization by

combining RankNet and third-party sources." In Proceedings of the EMNLP-CoNLL, pages 448–457.

[29] Barzilay, R. and Elhadad, M. (1997). "Using lexical chains for text summarization." in Proceedings ISTS'97.

[30] Hovy, E. and Lin, C. Y. (1999). "Automated text summarization in summarist." In Mani, I. and Maybury, M. T., editors, Advances in Automatic Text Summarization, pages 81–94. MIT Press

[31] N. Aletras and M. Stevenson. "Evaluating topic coherence using distributional semantics." In Proc. Of the 10th Int. Conf. on Computational Semantics (IWCS'13), pages 13–22, 2013.

[32] Kamal Sarkar "Automatic Single Document Text Summarization Using Key Concepts in Documents" J Inf Process Syst, Vol.9, No.4, pp.602-620, December 2013

[33] I. Chen "Integer Linear Programming Models for Constrained Clustering" in International Conference on Discovery Science 2010: Discovery Science pp 159-173

[34] Günes Erkan and Dragomir R. Radev. 2004. "LexRank: graph-based lexical centrality as salience in text summarization". J. Artif. Int. Res. 22, 1 (December 2004), 457-479.

[35] Jinqiang Bian, Zengru Jiang, Qian Chen 2014 "Research On Multi-document Summarization Based On LDA Topic Model" Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics 113-116

[36] Virendra Kumar Gupta Tanveer J. Siddiqui "Multi-Document Summarization Using Sentence Clustering" IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction, Kharagpur, India, December 27-29, 2012

[37] The Porter Stemming Algorithm [Online] Available:http://tartarus.org/~martin/PorterStemmer

[38] George A. Miller. "WordNet: A Lexical Database for English." Communications of the ACM, pages 39-41, November 1995

[39] Sherry and Dr. P. Bhatia " A Survey to Automatic Text Summarization Techniques" International Journal of Engineering Reasearch, October 2015 Pg. 1045- 1053

[40] Chin-Yew Lin and Eduard Hovy, "Identifying Topics by Position," In Proceedings of the Fifth conference on Applied natural language processing, San Francisco, pp. 283-290, 1997.

[41] S. P. Yong, A. I. Z. Abidin and Y. Y. Chen, "A Neural Based Text Summarization System," 6th International Conference of Data Mining, pp. 45-50, 2005.

[42] Ruqaiya Hasan, Coherence and Cohesive Harmony, In: Flood James (Ed.), Understanding Reading Comprehension: Cognition, Language and the Structure of Prose. Newark, Delaware: International Reading Association, pp. 181-219, 1984.

[43] William C. Mann and Sandra A. Thompson, Relational Propositions in Discourse, Defense Technical Information Center,

[44] Branimir Boguraev and Christopher Kennedy, "Saliencebased Content Characterization of Text Documents," In Proceedings of the ACL'97/EACL'97 Workshop on Intelligent Scalable Text Summarization, 1997.

[45] Li Chengcheng, "Automatic Text Summarization Based On Rhetorical Structure Theory," International Conference on Computer Application and System Modeling (ICCASM), vol. 13, pp. 595-598, October 2010.

[46] Xiaojun Wan, "An Exploration of Document Impact on Graph-Based Multi-Document Summarization," Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics,

[47] Tiedan Zhu and Xinxin Zhao, "An Improved Approach to Sentence Ordering For Multi-document Summarization," IACSIT Hong Kong Conferences, IACSIT Press, Singapore, vol. 25, pp. 29-33, 2012.

A Deep Structure of Person Re-Identification using Multi-Level Gaussian Models

Dinesh Kumar Vishwakarma, *IEEE Member*, Sakshi Upadhyay

***Abstract*—Person re-identification is being widely used in the forensic, and security and surveillance system, but person re-identification is a challenging task in real life scenario. Hence, in this work, a new feature descriptor model has been proposed using a multilayer framework of Gaussian distribution model on pixel features, which include color moments, color space values and Schmid filter responses. An image of a person usually consists of distinct body regions, usually with differentiable clothing followed by local colors and texture patterns. Thus, the image is evaluated locally by dividing the image into overlapping regions. Each region is further fragmented into a set of local Gaussians on small patches. A global Gaussian encodes, these local Gaussians for each region creating a multi-level structure. Hence, the global picture of a person is described by local level information present in it, which is often ignored. Also, we have analyzed the efficiency of earlier metric learning methods on this descriptor. The performance of the descriptor is evaluated on four public available challenging datasets and the highest accuracy achieved on these datasets are compared with similar state-of-the-arts, which demonstrate the superior performance.**

***Index Terms*—Feature Extraction, Feature Fusion, Image Matching, Image Recognition, Metric Learning, Person Re-Identification.**

I. Introduction

NOWADAYS, widespread networks of cameras are being used in various public places like railway stations, airports, hospitals, shopping malls, etc. which cover large areas and have non-overlapping viewpoints and thus provide a huge amount of relevant data. To utilize this data effectively for public safety applications we cannot rely on manual monitoring and thus need efficient automated systems, which can track person's activity across multiple cameras. Person re-identification (Re-Id) is an elementary task for this.

Person Re-Id aims to identify all the occurrences of a person of interest, in the data available through sensors (cameras), for a bounded region. It is mostly done by extracting some kind of visual property captured in the image. We assume that the data captured refers to a short duration of time so that the appearance of the person, specifically the clothes and the rare objects (like bag packs, purse, etc.), do not change. However, the difficulty level of this task is increased due to significant intra-class variations caused by the differences in illumination, pose and view angle of the camera. Moreover, due to inter-class similarity as a consequence of similar clothing and low resolution images acquired. This can be clearly seen in Fig.1 for some acquired public datasets. So the efforts are focused on developing discriminative features robust to these challenges and learning distance models for correct matching.

Most of the approaches followed in re-ID acquire appearance-based features [1] [2] [3] [4], such as color and texture models. To address the low resolution issue and varying poses, the color information like color histograms of desired color channels and color name descriptors [5] come to the rescue. However, they lack texture and spatial information and thus cannot sufficiently differentiate different persons of similar color. Thus texture descriptors, like Maximally Stable Color Regions (MSCR) [6], Local Binary Patterns (LBP) and 21 texture filters (8 Gabor filters and 13 Schmid filters), have been used in some works. However, it is seen earlier that texture descriptors do not give good results when used alone. Thus they are usually added as complimentary information to color and spatial features such that this fusion feature gives a rich quality signature. In [7], color histograms in two color spaces is combined with Gabor and Schmid filter responses and extracted on strips of predefined height and position. Maximally Stable Color Regions (MSCR) is fused with weighted Color Histograms (wHSV) in [8] [9], achieving outstanding results. Recently some high dimensional features, incorporating fusion, were also proposed. Some of them are SDALF [9], kBiCov



Fig. 1. Image pairs from VIPeR [4], QMUL GRID [29], PRID450s [30] and CUHK01 [31] datasets. Each column holds the images of the same individual captured from different cameras

[10], Fisher vectors (LDFV) [11], salience match [12], mid-level filter [13] and also ELF [4], where 8 color channels and 19 texture channels (Gabor and Schmid) were fused.

Methods for local feature encoding and pooling, as described in Fisher encoding [14] and Vector of locally Aggregated

D.K. Vishwakarma is working as an Associate Professor in the Department of Information Technology, Delhi Technological University, Delhi, India-11042 (e-mail: dinesh@dtu.ac.in, dvishwakarma@gmail.com)

S. Upadhyay is post graduate student in the Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India-110042 (e-mail: sakshiupadhyay8@gmail.com.)

Descriptors [15] (VLAD), are also the focus of many research works, since they have shown improvement on Bag of Words approach. One such popular feature with quite promising results is Covariance descriptor [16] [17] [18]. Earlier applied for matching and texture classification in [19] and then for pedestrian detection in [5], it estimated covariance of basic image features in overlapping regions. It gives a way to fuse multiple features naturally and also filters out noise samples during its computation, as it uses an average filter. To avoid overlapping of person a depth based video data [20] can be used but it results into a costly system and less field coverage in practical scenario. An image to video based person re-identification based on joint feature project matrix and heterogeneous dictionary pair learning [21] is developed to address the problem of variation within the videos.

In [21], the concept of shape of signal probability density function (SOSPDF) was presented which describes a signal based on its pdf, like a histogram and region covariance. However, an issue with using covariance is that it neglects the importance of mean information present in a local area. Therefore, it gave a descriptor named shape of Gaussians (SOG) as a new SOSPDF which inherited the properties of region covariance and also considered the mean information. Many other works also used such Gaussian models to improve the classification efficiency, like Huang and Wang [22] used it for face recognition. In [23] a Global Gaussian model was used to describe the probabilistic distribution of local features in an image. Some proposed to use Gaussian mixture model (GMM). In [24] hierarchical GMM was proposed in which the deviation of the distribution model of each image, with respect to the whole training set, was estimated instead of estimating GMM of each image, which would be impossible for a large scale data. However, it is not preferred due to its dependency on the training set. Another approach is Gaussian descriptor based on local features (GaLF) in which a 7-d feature including color, texture and spatial properties is extracted on each pixel and then the mean and covariance is calculated for region. However, it simply concatenates the two Gaussian parameters.

Inspired by these works and to overcome the drawbacks presented by them we have proposed a framework with multi-level encoding of pixel features, using Gaussian distribution. As we know that a human in an image has different body parts, with each part having some intra-similarity in features. So to utilize this local structure in a person image we have roughly divided the image into horizontal strips called regions. To create a deeper structure we have also worked on smaller consistent patches in the image with $k \times k$ pixel neighborhood. Then first we have estimated Gaussian for local patches and then for the entire region, by further extracting Gaussian parameters from the local patch Gaussians falling under that region. Former is called first level Gaussian (FG) and the latter is called second level Gaussian (SG). These region Gaussians are then concatenated to give the final descriptor. First level Gaussians are constructed on some kind of pixel features which can be color or texture data. For this, we have extracted color moments and Schmid filter responses in local patches, and Gradient information along with simple color values. Then we have created n-dimensional pixel features using different combinations of these basic pixel features and compared the performance.

Earlier a similar hierarchical model, covariance-of-covariance, has already been used effectively for image identification [25] in which region co-variance is approximated over local patch co-variances of pixel features. This is one of the sources of motivation for the use of two level distribution models in this work. However, the reason that it lacks the discriminative information provided by the mean data of pixels that's why we used the Gaussian model. Though, it handles the above issue but poses another challenge. Gaussian distribution lies on a particular Riemannian manifold, as per information geometry [26], where we cannot apply Euclidean functions while most existing discriminant metric methods just work in Euclidean space. Therefore, Riemannian manifold is projected onto a Euclidean space similar to the way described in [27] and [28], via. matrix logarithm and vectorization concept. Covariance-of-covariance approach [25] also used this type of solution to the Riemannian space problem.

In brief, the main contributions of the paper to the person re-identification are as follows:

- A deep structure is constructed using multilevel Gaussian model on four different combinations of colour, texture and spatial pixel features (colour moment and Schmid filter response).
- To make identification more effective a learning metric approach is applied on the developed structure and the identification accuracy is compared on four metric learning approaches.
- To evaluate the performance of the structure, an experiment is conducted on four challenging public datasets such as VIPeR [4], QMUL GRID [29], PRID450s [30] and CUHK01 [31]. The highest recognition rate is compared with the similar state-of-the-arts and demonstrate superior performance.
- In addition, to demonstrate the reliability of the structure, a person retrieval experiment is conducted and structure is able to re-identify the person.

The rest of paper is arranged as follows. The details of the proposed method is explained in section II, experimental and results are presented in section III and finally the conclusion of the work is given in section IV.

II. Multi-Level Gaussian Descriptor

We have constructed a deep structure of Gaussian model by extracting Gaussian distributions in two levels, one locally on the patch pixels, i.e. first level Gaussian and the other over the first level Gaussians in a defined region. The framework of our approach is shown in Fig 2. Motivation for this is taken from [25], [32], [33], [23] and [5]. It is named as a multi-level Gaussian descriptor. This structure is similar to some part based models described earlier where the image is divided into overlapping horizontal strips of fixed size given as one level of extraction and then each region is further dispersed into local

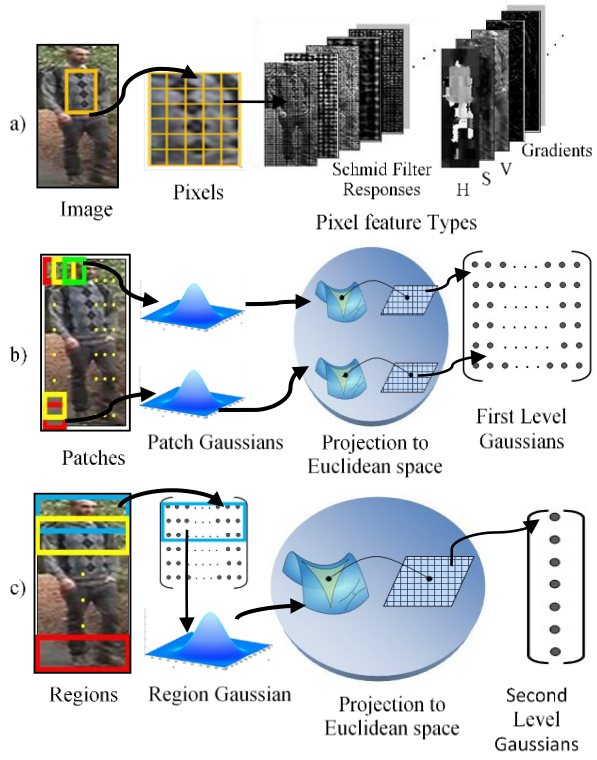


Fig. 2. Framework of the proposed Multi Level Gaussian Descriptor. It has three main sections (a) Extraction for low level pixel features (b) First level Gaussian encoding of patches (c) Second level Gaussian encoding of the regions over underlying patches.

square patches like the other level of extraction. Thus it addresses the non-rigid behavior of human body and captures the local information while representing the image globally.

The important aspect is to consider; the low-level pixel features, and it is necessary that the diverse aspects of the data contained in the image are extracted to add to the quality of base information on which further analysis is done. Thus in this method variety of base features are extracted and concatenated to create n-dimensional feature representation for each pixel before estimating the Gaussian model on them. The rest of the details are clearly given below.

A. *Low Level Features*

Define in this work, four different combinations of colour, texture and spatial pixel features are explored to create n-dimensional feature channel for each pixel in the image. Features used are color moment values of RGB components, HSV and normalized RGB color space values, Schmid filter responses, Gradient magnitudes in predefined orientations and spatial location of pixel.

Moments are statistical parameters that can uniquely describe a probabilistic distribution. Considering color in the image as a probabilistic distribution, moments on color values can provide a discriminative feature to measure the similarity between image pairs. They are inclusive of both shape and color data in the pixels and thus give an illumination, scale and rotation invariant feature, also used in [34] and [35]. In this work, first two moments of each R, G and B color values over local patches of $k \times k$ size, i.e. mean and standard deviation, are used in the feature channel. Color space values of HSV color space and normalized R and G components are the other color based features used in the feature channel. Color values are normalized, as described in [36], to reduce the effect of a change in saturation levels, illumination and color response of the camera sensors. This is done by dividing the color channel value by the summation of all channel values of that pixel.

Then Schmid filter banks are used to capture texture and edge information. They have rotational invariance property and can be estimated by convolving isotropic "Gabor-like" filters [37] with the image. They are orientation sensitive, similar to Gabor, and give circular gradients in the image. These filters combine frequency and scale together as follows:

$$F(r,\tau,\sigma) = F_0(\tau,\sigma) + \cos(\frac{2\pi\tau r}{\sigma})\exp(-\frac{r^2}{2\sigma^2}) \quad (1)$$

Here, r depicts the radius, τ depicts the number of cycles of the harmonic function within the Gaussian envelope of the filters, σ gives the scale and $F_0(\tau,\sigma)$ is a zero DC component which is added to obtain invariance to illumination changes. These filters are thus quite helpful in achieving invariance to intensity, pose and viewpoint.

Besides that, gradient information and spatial information of pixels are also added in the feature channel. Spatial location of a pixel in the vertical direction is used here because of the known symmetry present in the human body across the vertical axis, which is invariant to misalignments and viewpoint variations. The low-level features thus selected here have some useful properties to offer, but the selection is not too critical and finding better feature channels is very much a possibility.

B. *First Level Gaussian Encoding*

Once a low-level feature channel is extracted, image is transformed into a $W \times H \times n$ dimension representation where n denotes the dimensions of feature channel f_i for each pixel i and, $W \times H$ gives the image size. These pixel features are then encoded by a Gaussian distribution calculated over small patches of k x k pixels with u step size. For each patch P, the Gaussian model $\mathcal{N}(\mathfrak{f};\, \mu_p, \Sigma_p)$ is defined as:

$$\mathcal{N}(\mathfrak{f};\, \mu_p, \Sigma_p) = \frac{\exp(-\frac{1}{2}(\mathfrak{f}-\mu_p)^T \Sigma_{\mathrm{p}}^{-1}(\mathfrak{f}-\mu_p))}{(2\pi)^{n/2}|\Sigma_{\mathrm{p}}|} \quad (2)$$

Here μ_p is the mean vector and Σ_p is the covariance matrix estimated as $\mu_p = \frac{1}{N}\sum_{i\epsilon\mathcal{A}_p} f_i$ and $\Sigma_{\mathrm{p}} = \frac{1}{\mathrm{N}-1}\sum_{i\epsilon\mathcal{A}_p}(f_i - \mu_p)(f_i - \mu_p)^{\mathrm{T}}$, respectively. $\mathcal{A}_p$ denotes the area of sampled patch P and N gives the number of pixels in $\mathcal{A}_p$. $\mathfrak{f}, \mu_p \epsilon \mathbb{R}^n$ $and \Sigma_{\mathrm{p}} \epsilon \mathbb{S}_{\mathrm{n}}^{+}$ where $\mathbb{S}_{\mathrm{n}}^{+}$ is the space of real symmetric positive semi-definite (SPD) matrices. Estimated covariance matrix is $n \times n$ dimensional with diagonal entries representing the variance of each feature, and the non-diagonal entries giving their respective correlations.

So, in this step, the first level of Gaussian encoding is extracted over the pixel features. But for the second level of Gaussian encoding, we need to evaluate two mathematical parameters, i.e. mean and covariance, on a set of first level Gaussians. As per information geometry [26], Gaussians belong

to Riemannian manifold where Euclidean functions cannot be applied. So it needs to be projected from the Riemannian space to a tangent Euclidean space favored by Riemannian metric, as described in [5] [38] [39] [27]. This mapping is performed over a Symmetric positive definite matrix. So firstly, the n-dimensional Gaussian distributions are embedded into a n+1 dimensional SPD matrix space $\mathbb{S}_{n+1}^{+}$, similar to the work in [28], as follows:

$$\mathcal{N}(\mathfrak{f};\, \mu_p, \Sigma_p) \sim S_p = |\Sigma_p|^{-1/(n+1)} \begin{bmatrix} \Sigma_p + \mu_p \mu_p^T & \mu_p \\ \mu_p^T & 1 \end{bmatrix}, \quad (3)$$

where S_p is the$(n+1) \times (n+1)$ dimensional SPD matrix derived from the local Gaussians. This is then mapped onto a space tangent to the manifold, at the tangency point given by matrix T, via simple matrix logarithm, as defined in [5]:

$$s = log_T(S_p) \triangleq T^{\frac{1}{2}} \log(T^{-\frac{1}{2}} S_p T^{-\frac{1}{2}}) T^{\frac{1}{2}} \quad (4)$$

Here, s is the tangent vector in the tangent space at point T, which is further minimized by vectorization, given as:

$$S_p' = vec_T(s) = vec_I(T^{-\frac{1}{2}} s T^{-\frac{1}{2}}) \quad (5)$$

S_p' denotes the orthonormal coordinates of S which lie in a Euclidean space and I denotes an identity matrix. Replacing the value of 'S ' in (5) by the value defined in (4) redefines S_p', the projection vector of S_p, as follows:

$$S_p' = vec_I(\log(T^{-\frac{1}{2}} S_p T^{-\frac{1}{2}})) \quad (6)$$

Therefore, the matrix of patch Gaussian S_p becomes $m = \frac{(n^2+3n)}{2} + 1$ dimensional vector S_p'.

C. Second Level Gaussian Encoding

Now the second level of encryption is done over the patch Gaussians obtained previously. Each image is roughly divided into seven overlapping horizontal stripes, called regions, to utilize the local geometry in humans. Then these regions are again encoded by a Gaussian distribution $\mathcal{N}(S'; \mu^R, \Sigma^R)$ estimated over the first level Gaussians falling under the region R. Thus a discriminative and compact model of Gaussian distributions is learnt. But to use this feature for classification and matching purpose, again we need to apply SPD matrix conversion and projection to Euclidean space, as we did previously. This is required because most of the existing metric methods work in Euclidean space. Then the SPD matrix obtained will have $(m+1) \times (m+1)$ dimensions and the projected vector $\boldsymbol{H}$ will have $(m^2 + 3m/2) + 1$ dimensions.

The final feature vector is then formed by just concatenating all the region feature vectors $\{\boldsymbol{H}_r\}_{r=1}^{R}$ into a single feature vector $\boldsymbol{H} = [\boldsymbol{H_1^T}, \ldots, \boldsymbol{H_R^T}]^T$. Due to deep structure of the model final feature vector is very high dimensional and thus L2 norm normalization is performed. But prior to this, mean removal is done on the feature vectors to deal with the large differences in the values at different dimensions, as described in [40]. This may occur due to the use of variety of pixel features with different distribution characteristics. Then the normalization is defined as: $\boldsymbol{H} = (\boldsymbol{H} - \overline{\boldsymbol{H}})/\| \boldsymbol{H} - \overline{\boldsymbol{H}} \|_2$ where $\overline{\boldsymbol{H}}$ is the mean vector of the training set feature vectors.

D. Metric Learning Methods

Once the descriptor is constructed then, we need to compare the derived features to find the most similar resemblances. For this, robust and efficient distance metrics must be used. They must try to minimize the distance values between the same class objects while maximizing those from different classes. In this work, four metric learning methods are compared against the proposed descriptor. These are SVMML [41], LFDA [42], KISSME [43] and XQDA [32].

SVMML, proposed in [41], learns a decision function locally by using a distance metric and a locally adaptive threshold rule jointly, following SVM-like objective function approach. LFDA proposed in [42] learns a distance metric using a two stage processing, where first PCA is applied for dimensionality reduction and then supervised FDA is used along with LPP for further reduction. Then KISSME was proposed in [43] as an effective metric just based on equivalence constraints. It does not involve tedious iterative computation for optimizing the procedure. It just needs to compute two small sized co-variance matrices and is thus scalable to large datasets. An extension of KISSME, LFDA and Bayesian Metric methods was then proposed in LOMO [32], named as Cross view discriminant analysis or XQDA. The statistical inference perspective based on likelihood-ratio test as given in KISSME also inspires it. It can learn a lower dimension discriminant subspace and a metric kernel on it, simultaneously, by following generalised Rayleigh-quotient formulation. While LFDA, after deriving a subspace directly uses the Euclidean function on it, XQDA further approach to use an efficient metric as well.

III. Experiments and Results

To show the efficiency of our method we have performed various experimental studies, including comparisons with state-of-the-art descriptors and metrics, on four challenging datasets: VIPeR [4], QMUL GRID [29], PRID450s [30] and CUHK01 [31]. The results are presented in the form of Cumulative Matching Characteristics (CMC) curves and Recognition rates at different ranks. We have also performed image retrieval experiment on some of the datasets considering person re-identification as a recognition problem.

A. Datasets and Setup

VIPeR [4] has a total of 1264 images of 632 individuals captured from two different camera views, A and B. The images have illumination and viewpoint variations, and are cropped and resized to 128 × 48 pixels. For evaluation, the widely adopted protocol is used where the set of 632 image pairs is randomly split into training, and a testing set with 316 image pairs in each and 10 such sets are created to produce average results.

QMUL GRID [29] contains 250 low-resolution pedestrian image pairs captured in a busy underground station by two cameras and contains variation in poses and lighting. Two folders are provided namely 'Probe' and 'Gallery' each containing 250 images of 250 individuals while gallery folder also contains additional 775 images that do not belong to probe set and are used in the testing set during cross validation.

PRID450S [30] contains 450 images captured from disjoint views by two cameras A and B. For evaluation, 225 individuals are assigned randomly to training and testing set.

CUHK01 [31] has 3884 images in a total of 971 individuals manually cropped to 60x160 pixels. Two disjoint camera view were selected where camera A captures more pose and viewpoint variations while camera B captures mainly frontal view and back view. Each person has two images captured in each camera. The training set contains 486 while testing set contains 485 individuals selected randomly.

Each image in all these datasets is resized to 128×48 pixels. For each image pair, image from one camera is assigned as a probe and the other camera as a gallery, randomly. It is done in both training and testing phase. Then the process of selecting a single image from the probe set and matching it with all images from the gallery set is repeated for all images in the probe set.

B. Feature settings

We have used four different sets of n-dimensional pixel level feature vector.

1) Pixel Feature 1 (YCM)

In this y distance and first two color moments, i.e. mean and standard deviation of RGB color space are used to create a 7-d pixel feature. Color moments are calculated on patches of size 5×5 with patch interval of size 2. It can be represented as:

$$[y, mean(R), std(R), mean(G), std(G), mean(B), std(B)]$$

2) Pixel Feature 2 (SCHMID)

This is a texture feature representation of pixels comprising of 13-d schmid filter responses using schmid filter banks applied on 10×10 non-overlapping patches. Parameter values of (σ, τ) for 13 filter banks are (2,1), (4,1), (4,2), (6,1), (6,2), (6,3), (8,1), (8,2), (8,3), (10,1), (10,2), (10,3) and (10,4).

3) Pixel Feature 3 (YGOHSV

The third kind of feature combination used has 1-d y distance value, 4-d gradient orientation and 3-d HSV color space values, i.e. a 8-d pixel feature representation.

4) Pixel Feature 4 (YGOnRnG)

It is a 7-d pixel feature representation comprising of 1-d y distance value, 4-d gradient orientation and 2-d nRnG color space values. Here the nRGB is the normalized color space obtained by normalising RGB color space values (e.g., nR = R/(R+G+B)). We have used only {nR, nG} values, since this color space has redundancy.

After constructing n-dimensional pixel feature we have extracted multi-level Gaussian descriptor for seven horizontal strips or regions (R=7) that are overlapping. Each strip comprises 32×48 pixels. In each region local patch Gaussians are extracted from the patches of 5×5 pixels with patch interval 2.

C. Performance Comparison of Different Pixel Features

The effectiveness of each pixel feature and their fusion is evaluated separately, and the results are shown by CMC curves in Fig. 3, Fig. 4 and Fig. 5, for VIPeR, PRID450s and QMUL GRID, respectively. The identification rates at rank-1, rank-10 and rank-20, for different combinations of pixel features, are given in Table I. XQDA metric is used for this evaluation. Rank-1 recognition rate of YCM pixel feature is highly greater than the pixel feature comprising of Schmid filter responses, for VIPeR and PRID450s. It proves that for person re-id problem,

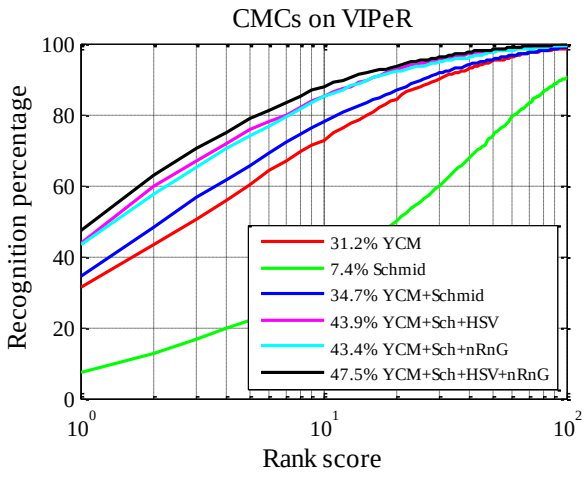


Fig. 3. CMC curve and Rank-1 identification rates on VIPeR

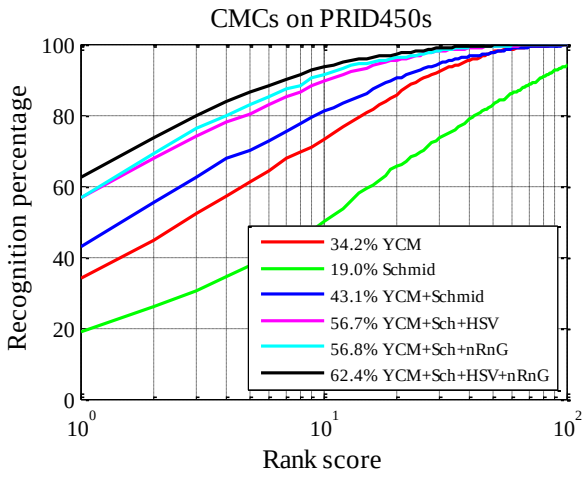


Fig. 4. CMC curve and Rank-1 identification rates on PRID450s

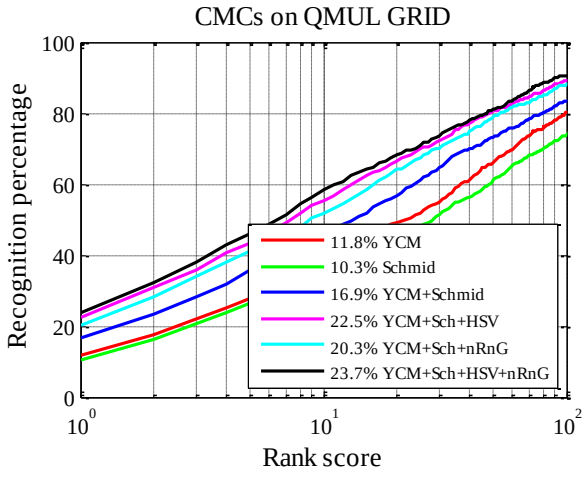


Fig. 5. CMC curve and Rank-1 identification rates on PRID450s

TABLE I
COMPARISON OF RECOGNITION RATES FOR DIFFERENT COMBINATION OF PIXEL FEATURES ON VIPER, PRID450S AND GRID DATASETS. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Dataset	Pixel Feature Used	R=1 (%)	R=10 (%)	R=20 (%)
VIPeR	YCM	31.2	72.9	86.8
	Schmiid	7.4	34.2	50.1
	YCM+Schmid	34.7	78.1	86.8
	YCM+Schmid+HSV	43.9	85.1	92.5
	YCM+Schmid+nRnG	43.4	85.1	92.0
	YCM+Schmid+HSV+nRnG	**47.5**	**87.9**	**93.7**
PRID450s	YCM	34.2	73.3	85.7
	Schmiid	19.0	49.8	65.3
	YCM+Schmid	43.1	81.1	90.3
	YCM+Schmid+HSV	56.7	89.6	95.4
	YCM+Schmid+nRnG	56.8	91.2	96.0
	YCM+Schmid+HSV+nRnG	**62.4**	**93.5**	**96.9**
QMUL GRID	YCM	11.8	38.2	49.2
	Schmiid	10.3	35.4	49.2
	YCM+Schmid	16.9	45.7	56.7
	YCM+Schmid+HSV	22.5	55.3	66.4
	YCM+Schmid+nRnG	20.3	52.0	64.2
	YCM+Schmid+HSV+nRnG	**23.7**	**58.2**	**68.1**

*YCM+Schmid+HSV+nRnG = proposed descriptor (MLGD).
Evaluation performed using XQDA metric.

color features perform much better than texture features. But texture feature when combined with color feature, exceeds the rank-1 recognition rates of those using the color feature alone. The best results are achieved when all the four feature types are fused to make a rich signature. We have proposed this fusion descriptor as Multi-Level Gaussian descriptor (MLGD).

D. Performance Comparison of the Proposed Descriptor on Different Metrics

The proposed descriptor MLGD is evaluated by four different metric learning methods namely, SVMML [41], LFDA [42], KISSME [43] and XQDA [32]. The results on VIPeR dataset are given in Fig. 6 by CMC curves and the recognition rates at rank-1, 10 and 20 are given in Table II. It is clearly visible that the proposed descriptor gives good rank-1 recognition rates, greater than > 35%, when evaluated with these state-of-the-art metric learning methods. Thus it can be inferred that the proposed descriptor holds a robust structure and is an effective representation for images. Best results, with rank-1 recognition rates of 47.53%, are achieved by using XQDA metric.

E. Performance Comparison with State-of-the-Art Methods

The proposed descriptor is compared with both the single as well multi-level descriptors and the performance is listed in Table III. Analysis on covariance distribution models show that the Covariance descriptor when used globally on the image as proposed in [19], gives 26.9% rank-1 rates on VIPeR dataset while a multi-level model of covariance, proposed as Cov-of-Cov in [44] [25], gives 33.9% rank-1 rates on VIPeR. It is approximately 7% better than the global model of covariance.

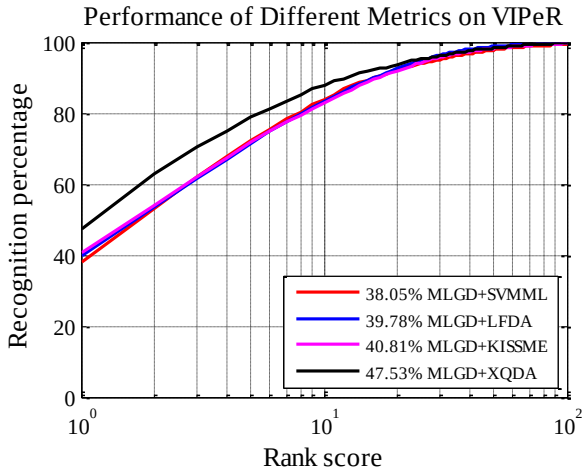


Fig. 6. CMC curve and Rank-1 identification rates of the descriptor evaluated on four different metrics.

TABLE II
COMPARISON OF RECOGNITION RATES FOR STATE-OF-THE-ART METRICS APPLIED ON THE PROPOSED DESCRIPTOR. BEST RESULTS ARE HIGHLIGHTED IN BOLD

Metrtic Methods	Rank=1 (%)	Rank=10 (%)	Rank=20 (%)
MLGD+SVMML [41]	38.05	83.90	92.02
MLGD+LFDA [42]	39.78	83.21	92.59
MLGD+KISSME [43]	40.81	82.80	91.85
MLGD+XQDA [32]	**47.53**	**87.92**	**93.71**

There are number of single-level meta descriptors such as Heterogeneous Auto-Similarities of Characteristics (HASC) [45], Hybrid Spatiogram and Covariance Descriptor (HSCD) [16], Local Descriptors encoded by Fisher Vector (LDFV) [11], Second-order Average Pooling (2AvgP) [46], and GOLD [47] are also chosen for comparison of the performance listed in Table III. They are non-hierarchical models and ignore the local properties of regions. They have similar performances. It can be seen that our descriptor has the highest recognition rates, with the rank-1 rate of 47.5% on VIPeR. It is also evident from the table that both the hierarchical methods, i.e. our method and Cov-of-Cov [25], have higher rank-1 matching rates than the non-hierarchical ones and thus it can be inferred that hierarchical model usually performs better. Between the two hierarchical methods mentioned, the difference is that we have used mean information together with the covariance value while [25] has used just the covariance information. Since our method has 13.6 % better rank-1 rates than hierarchical covariance descriptors so this proves that means information is a valuable feature of the image.

In section (c) of Table III, we have listed the reported results on some of the state-of-the-art method, including Metric Ensemble [48], Mid-Level Filter Learning (MLFL) [13], Salience Matching [12] and LOMO [32]. It can be clearly noted from the results that the performance of MLGD goes above many state-of-the-art methods with 47.5%, 62.4%, 54.5% and 23.7% rank-1 rates on VIPeR, PRID450S, CUHK01 and GRID dataset, respectively. The proposed descriptor even outperforms the efficient metric ensemble [48] by 1.6 % rank-1 rates.

TABLE III.
PERFORMANCE COMPARISON OF THE PROPOSED DESCRIPTOR MLGD AND STATE-OF-THE-ART DESCRIPTORS ON VIPER, PRID450S, QMUL GRID AND CUHK01. SECTION (A) AND (B)REPRESENTS MULTI LEVEL DESCRIPTORS AND SINGLE LEVEL DESCRIPTORS, RESPECTIVELY, EVALUATED USING XQDA METRIC. SECTION (C) SHOWS THE PERFORMANCE OF SOME STATE-OF-THE-ART METHODS. THE BEST RESULTS IN EACH SECTION ARE GIVEN IN RED AND THE SECOND BEST IN BLUE.

Section	Descriptors	VIPeR			PRID450s			GRID			CUHK01		
		R=1 (%)	R=10 (%)	R=20 (%)	R=1 (%)	R=10 (%)	R=20 (%)	R=1 (%)	R=10 (%)	R=20 (%)	R=1 (%)	R=10 (%)	R=20 (%)
(A)	MLGD(Our method)	**47.5**	87.9	**93.7**	**62.4**	**93.5**	**96.9**	**23.7**	**58.2**	**68.1**	**54.5**	**83.5**	**90.5**
	COV-of-COV [25]	33.9	76.6	87.7	47.0	83.4	91.6	16.6	45.0	55.2	40.9	72.5	81.1
(B)	Cov [19]	26.9	65.8	77.1	40.4	73.4	82.1	10.6	29.0	36.7	34.5	64.5	73.6
	GOLD [47]	27.1	66.5	77.7	40.5	73.8	82.2	10.9	29.2	37.4	35.3	65.2	74.2
	HSCD [16]	31.2	86.5	91.8	-	-	-	-	-	-	-	-	-
	2AvgP [46]	28.8	68.5	79.2	44.7	75.8	83.8	12.9	36.7	47.4	36.1	68.1	76.3
	LDFV [11]	25.3	66.8	79.4	32.1	66.9	77.6	16.2	41.9	53.1	36.4	71.0	80.3
	HASC [45]	30.9	70.6	81.8	41.8	76.3	85.2	12.9	35.6	47.3	38.6	68.7	77.1
	CH+LBP [49]	27.7	69.3	82.4	21.5	60.8	74.4	16.2	45.0	57.1	31.3	70.4	81.5
(C)	Metric Ensemble [48]	45.9	88.9	95.8	-	-	-	-	-	-	53.4	84.4	90.5
	LOMO [32]	41.1	82.2	91.1	62.6	92.0	96.6	17.9	46.3	56.2	49.2	84.2	90.8
	SalMatch [12]	30.2	65.0	-	-	-	-	-	-	-	28.5	55.0	-
	MLFL [13]	29.1	65.9	70.9	-	-	-	-	-	-	34.3	65.0	75.2

F. Image Retrieval Results

We have also performed an image retrieval experiment where we have tried to retrieve the true match of a probe image from the disjoint camera images. We have performed this on three datasets namely, VIPeR, CUHK01and PRID450s and displayed the results in the Fig. 7, Fig. 8 and Fig. 9.



Fig. 7. Image retrieval results on VIPeR. Yellow boxes represent the query image and Red boxes give the true match pair of the query image.



Fig. 8. Image retrieval results on PRID450s Yellow boxes represent the query image and Red boxes give the true match pair of the query image.



Fig. 9. Image retrieval results on CUHK01. Yellow boxes represent the query image and Red boxes give the true match pair of the query image.

We have first given a probe image to the system and then retrieved closest 10 images from the disjoint gallery set. Out of these, one of the 10th retrieved images contains the true match of the probe image from a different view angle and is highlighted by Red boxes in the given figures. The results clearly show that the recognition accuracy of the system is quite high, as the closest matches found are also the true match pair of the probe image. It must be mentioned here that while VIPeR and PRID450s have single frame per person in each camera images, CUHK01 has two or more image frames per person in each camera images.

IV. CONCLUSION

This work presents an efficient and effective approach of multi-layer Gaussian descriptor model for the problem of person re-ID. The descriptor presented in this work utilizes both, mean and covariance values of pixel features, present in an image and thus returns a robust and discriminative representation of data. It first encodes the local patches by a Gaussian distribution and then performs a second level of encoding over a set of local patches falling under overlapping regions, again by a Gaussian model. In this way it does not

neglect the relevant information in the local structures of the image while acquiring a global representation. The results of our in depth experiments proved that the proposed descriptor can even outperform the state-of-the-art performances on four public datasets. We have also conducted image retrieval experiment where we treat this problem as a recognition problem. The results prove that our descriptor is robust against viewpoint changes and illumination variations. Further we have analyzed four different efficient metrics on our descriptor to show its robust structure.

In future the deep structure of Gaussian can be explored further to consider the local structure of human appearances in more depth. There is ample of scope in finding computationally low cost features to reduce the extraction and search time of the system.

In addition, some other kinds of pixel features can be tested to search for any further improvements in the process accuracies since the selection of these features is not too critical and finding better feature channels is very much a possibility. Then we can also try to learn a new metric, which is highly discriminative for high dimensional feature vectors.

V. REFERENCES

[1] W. Zhang, B. Ma, K. Liu and R. Huang, "Video-Based Pedestrian Re-Identification by Adaptive Spatio-Temporal Appearance Model," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 2042-2054, 2017.

[2] M. Farenzena, L. bazzani, A. Perina, V. Murino and M. Critani, "Person re-identification by symmetry-driven accumulation of local features," in *IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010.

[3] W. Lin , Y. Shen, J. Yan , M. Xu, J. Wu, J. Wang and K. Lu, "Learning Correspondence Structures for Person Re-Identification," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2438-2453, 2017.

[4] D. Gray and H. Tao, "Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features,," in *ECCV*, 2008.

[5] O. Tuzel, F. Porikli and P. Meer, "Pedestrian Detection via Classification on Riemannian Manifolds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1713-1727, 2008.

[6] P. E. Forssen, "Maximally Stable Colour Regions for Recognition and Matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, Minnepolis, 2007.

[7] M. Dikemen, E. Akbas, T. S. Huang and N. Ahuja, "Pedestrian Recognition with a Learned Metric," in *Asian Conference on Computer vision (ACCV)*, Queenstown, 2011.

[8] B. Alipanahi, M. Biggs and A. Ghodsi, "Distance metric learning vs. fisher discriminant analysis," in *Proceedings of the 23rd national conference on Artificial intelligence*, 2008.

[9] L. Bazzani, V. Murino and M. Cristani, "Symmetry-driven accumulation of local features for human characterization and re-identification," *Computer Vision and Image Understanding*, vol. 117, no. 2, pp. 130-144, 2013.

[10] B. Ma, Y. Su and F. Jurie, "Covariance descriptor based on bio-inspired features for person re-identification and face verification," *Image and Vision Computing*, vol. 32, no. 6, pp. 379-390, 2014.

[11] B. Ma, Y. Su and F. Jurie, "Local descriptors encoded by fisher vectors for person re-identification," in *European Conference on Computer Vision (ECCV)*, Florence, 2012.

[12] R. Zhao, W. Ouyang and X. Wang, "Person re-identification by salience matching," in *International Conference on Computer Vision*, 2013.

[13] R. Zhao, W. Ouyang and X. Wang, "Learning Mid-level Filters for Person Re-identification," in *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 2014.

[14] F. Perronnin, J. S´anchez and T. Mensink, "Improving the fisher kernel for large-scale image classification," in *ECCV'10 Proceedings of the 11th European conference on Computer vision*, Heraklion, Crete, Greece, 2010.

[15] H. Jegou, M. Douze, C. Schmid and P. Perez, "Aggregating local descriptors into a compact image representation," in *CVPR*, San Francisco, 2010.

[16] M. Zeng, Z. Wu, C. Tian, L. Zhang and L. Hu, "Efficient person re-identification by hybrid spatiogram and covariance descriptor," in *CVPR Workshops*, Boston, 2015.

[17] S. Bak, G. Charpiat, E. Corv´ee, F. Br´emond and M. Thonnat, "Learning to match appearances by correlations in a covariance metric space.," in *European Conference on ComputerVision (ECCV)*, Florence, 2012.

[18] S. Bak, E. Corv´ee, F. Br´emond and M. Thonnat, "Boosted human re-identification using Riemannian manifolds," *Image Vision and Computing*, vol. 30, no. 6, pp. 443-452, 2012.

[19] O. Tuzel, F. Porikli and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *ECCV*, 2006.

[20] A. Wu, W. S. Zheng and J. H. Lai, "Robust Depth-Based Person Re-Identification," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2588-2603, 2017.

[21] X. Zhu, X. Y. Jing, X. You, W. Zuo, S. Shan and W. S. Zheng, "Image to Video Person Re-Identification by Learning Heterogeneous Dictionary Pair With Feature Projection Matrix," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 717-732, 2018.

[22] L. Gong, T. Wang and F. Liu, "Shape of Gaussians as feature descripors," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[23] Z. Huang, R. Wang, S. Shan, X. Li and X. Chen, "Log-Euclidean metric learning on symmetric positive definite manifold with application to image set classification.," in *International Conference on Machine Learning (ICML)*, Lille, 2015.

[24] H. Nakayama, T. Harada and Y. Kuniyoshi, "Global Gaussian approach for scene categorization using information geometry," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, 2010.

[25] X. Zhou, N. Cui, Z. Li, F. Liang and T. S. Huang, "Hierarchical Gaussianization for image classification," in *12th IEEE International Conference on Computer Vision (ICCV)*, Kyoto, 2009.

[26] G. Serra, C. Grana, M. Manfredi and R. Cucchiara, "Covariance of covariance features for image classification," in *Proceedings of International Conference on Multimedia Retrieval*, Glasgow, 2014.

[27] S. I. Amari, "Information geometry on hierarchy of probability distributions," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1701-1711, 2001.

[28] V. Arsigny, P. Fillard, X. Pennec and n. Ayache, "Geometric means in a novel vector space structure on symmetric positive-definite matrices," *SIAM J. Matrix Analysis Applications*, vol. 29, no. 1, p. 328–347, 2006.

[29] P. Li, Q. Wang and L. Zhang, "A Novel Earth Mover's Distance Methodology for Image Matching with Gaussian Mixture Models," in *ECCV*, Sydney, 2013.

[30] C. C. Loy, T. Xiang and S. Gong, "Time-delayed correlation analysis for multi-camera activity understanding," *International Journal of Computer Vision*, vol. 90, no. 1, pp. 106-129, 2010.

[31] P. M. Roth, M. Hirzer, M. K¨ostinger, C. Beleznai and H. Bischof, "Mahalanobis distance learning for person reidentification," in *Person Re-Identification*, 2014, pp. 247-267.

[32] W. Li, R. Zhao and X. Wang, "Human reidentification with transferred metric learning," in *Asian Conference on Computer Vision*, 2012.

[33] S. Liao, Y. Hu, X. Zhu and S. Z. Li, "Person re-identification by local maximal occurrence representation and metric learning," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 2015.

[34] B. Ma, Q. Li and H. Chang, "Gaussian descriptor based on local features for person re-identification," in *Asian Conference on Computer Vision*, Singapore, 2014.

[35] Z. C. Huang, P. Chan, W. W. Y. Ng and D. S. Yeung, "Content-based image retrieval using color moment and Gabor texture feature," in *ICMLC*, Qingdao, 2010.

[36] H. Yu, M. Li, H. Zhang and J. Feng, "Color texture moments for content-based image retrieval," in *ICIP*, 2002.

[37] Y. Du, H. Ai and S. Lao, "Evaluation of color spaces for person reidentification," in *Proceedings of the 21st International Conference on Pattern Recognition*, Washington, 2012.

[38] C. Schmid, "Constructing models for content-based image retrieval," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

[39] V. Arsigny, P. Fillard and N. Ayache, "Fast and Simple Computations on Tensors with Log-Euclidean Metrics," [Research Report] RR-5584, INRIA, 2005.

[40] X. Pennec, P. Fillard and N. Ayache, "A Riemannian framework for tensor computing," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41-66, 2006.

[41] H. Jegou and O. Chum, "Negative evidences and co- occurences in image retrieval: The benefit of PCA and whitening," in *ECCV*, 2012.

[42] Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao and J. Smith, "Learning Locally-Adaptive Decision Functions for Person Verification," in *IEEE Conference on Computer Vision and Pattern Recognition*, Portland, 2013.

[43] S. Pedagdi, J. Orwell, S. Velastin and B. Boghossian, "Local Fisher Discriminant Analysis for Pedestrian Re-identification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013.

[44] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth and H. Bischof, "Large scale metric learning from equivalence constraints," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[45] Q. Wang and P. Li, "Local log-Euclidean covariance matrix (L2ECM) for image representation and its applications.," in *European Conference on Computer Vision (ECCV),*, 2012.

[46] M. S. Biagio, M. Crocco, M. Cristani, S. Martelli and V. Murino, "Heterogeneous auto-similarities of characteristics (HASC): exploiting relational information for classification," in *IEEE International Conference on Computer Vision*, 2013.

[47] J. Carreira, R. Caseiro, J. Batista and C. Sminchisescu, "Free-Form Region Description with Second-Order Pooling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 6, pp. 1177-1189, 2015.

[48] G. Serra, C. Grana, M. Manfredi and R. Cucchiara, "GOLD: Gaussian of local descriptors for image representation," *Computer Vision and Image Understanding*, vol. 134, pp. 22-32, 2015.

[49] S. Paisitkriangkrai, C. Shen and A. V. D. Hengel, "Learning to rank in person re-identification with metric ensembles," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[50] F. Xiong, M. Gou, O. Camps and M. Sznaier, "Person re-identification using kerne based metric learning methods," in *Europian conference on computer vision (ECCV)*, 2014.



Dinesh Kumar Vishwakarma 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, an d the Ph.D. degree in computer vision from Delhi Technological University, New Delhi, India, in 2016. He is currently an Associate Professor with the Department of Information Technology, Delhi Technological University, Delhi, India. His current research interests include human action and activity recognition, hand gesture recognition, gait analysis, and machine learning.

Dr. Vishwakarma is a Reviewer of various journals of IEEE/IET, Springer, and Elsevier.



Sakshi Upadhyay Received B.Tech degree in Electronics and Communication Engineering from BT Kumaon Institute of Technology, Dwarahat, Uttarakhand, India in 2014, and M.Tech. in Signal Processing and Digital Design from Delhi Technological University, New Delhi, in 2017. Her research interests include Person Identification, Digital Image Processing, Machine learning and Pattern Recognition.

Journal of Information Sciences and Computing Technologies(JISCT)
ISSN: 2394-9066

SCITECH
RESEARCH ORGANISATION|

Volume 7, Issue 2
April 28, 2018

Journal of Information Sciences and Computing Technologies
www.scitecresearch.com/journals

A Hand-Based Biometric Verification System Using Ant Colony Optimization

Nirmal Pandey[1], Prof O.P. Verma[2], Dr Amioy Kumar[3]
[1]Department of Computer Science & Engineering
Delhi Technological University
New Delhi – 110042
[2]Principal G B Pant Engineering College,
Govt. Of NCT of Delhi, Okhla
New Delhi – 110020
[3]Formerly Guest Faculty (Department of Computer Science & Engineering)
Delhi Technological University
New Delhi – 110042

Abstract

This paper presents a novel personal authentication system using hand-based biometrics, which utilizes internal (beneath the skin) structure of veins on the dorsal part of the hand and the outer shape of the hand. The hand-vein and the hand-shape images can be simultaneously acquired by using infrared thermal and digital camera respectively. A claimed identity is authenticated by integrating these two traits based on the score-level fusion in which four fusion rules are used *for the integration.* Before their fusion, each modality is evaluated individually in terms of error rates and weights are assigned according to their performance. In order to achieve an adaptive security in the proposed bimodal system, an optimal selection of fusion parameters is required. Hence, Ant Colony Optimization (ACO) is employed in the bimodal system to select the weights and also one out of the four fusion rules optimally for the adaptive fusion of the two modalities to meet the user defined security levels. The databases of hand-veins and the hand-shapes consisting of 150 users are acquired using the peg-free imaging setup. The experimental results show genuine acceptance rate (GAR) of 98% at false acceptance rate (FAR) of 0.001% and the system has the potential for any online personal authentication based application.

***Index Terms*:** Biometric Authentication, Infrared thermal vein patterns, hand-shapes, Score-level fusion, ACO.

I. INTRODUCTION

Biometric based personal authentication is aimed at determining the claimed identity based on user's physiological/behavioral traits [1]. The biometric authentication system utilizes stepwise procedures. In the first step, a sensor acquires the biometric modality which is converted into the digital image. In the second step, the feature extraction module extracts the reliable features from the digital image and a data storage component saves the biometric feature templates corresponding to each enrolled user. In the third step, the matching module compares the test biometric template with the training templates and provides both the genuine and imposter matching scores. In the fourth step, the decision module acts upon the matching scores to learn a decision threshold between the genuine (acceptance) and the imposter (rejection) for the claimed

identity [1]. A conventional way to choose the best decision-threshold to improve the performance is via Receiver Operating Characteristics Curve (ROC) which is a plot of false acceptance rate (FAR) vs. false rejection rate (FRR) [1]. The mode of authentication in a biometric system can be categorized into two types: identification and verification. Identification is the 1: N match obtained by matching the test pattern with each of the enrolled patterns. The unknown identity can be traced out by providing the best match among *N* matches [1]. The verification is a 1:1 whereby the claimed identity of a pattern is verified by matching it with the enrolled patterns. If the user identity is unknown, the authentication can only be performed under the identification mode [11]-[14]. When the user's identity is known, verification is the most preferred mode for the personal authentication [6]-[10]. This paper presents a novel hand based bimodal system using hand veins and hand shape for biometric based verification. It is designed to adaptively select the fusion parameters according to the desired security requirements.

A. Hand-Based Bimodal System

Hand-based biometrics, particularly palmer part has been studied extensively in the recent years and many authentication systems are developed using fingerprint [2]-[4], palmprint [5]-[7] and hand-geometry [8]-[9] with high user acceptance. However, being the outer most part of the human body people leave their fingerprints and palm impressions unconsciously. This sort of negligence is prone to security risks on these modalities.

The slowly gaining popularity of the vein patterns as a promising biometric trait for the personal authentication is due to its ability to defy imposter attacks. The veins being the blood carrying vessels are a part of the vascular system of the body and any change in the body generally affects the vascular system. Normally, it expands with the growth of the body from the childhood and remains fairly stable between 16 to 60 years of age. The vein patterns of persons lying within this age group can serve as a unique biometric. Located beneath the skin, vein patterns are difficult to spoof by an imposter attack, and hence are a safer biometric trait. But then, being a hidden structure of the human body, the acquisition of vein patterns are highly susceptible to environmental conditions like ambient temperature and humidity [10]. Therefore, the need is to make a bimodal system where the vein patterns can be integrated with other biometric traits.

In this work, a bimodal system is developed using vein patterns extracted from the dorsal part of the hand and the contour information from the same part of the hand, called as *hand-shape*. The infrared thermal camera is used to capture hand-vein patterns while visible light camera is used to acquire hand-shape images. Both the biometric traits can be simultaneously extracted from the claimed identity and integrated before final accept/reject decision. The objectives of this bimodal system are: (i) To improve the overall performance, (ii) to increase the system accuracy, and (iii) to withstand imposter attacks whenever deployed for the personal authentication.

B. The Bimodal System Using ACO

The widely used strategies for the fusion of two modalities in a bimodal systems are score-level fusion [11] and the decision-level fusion [11]. A typical multi biometric system requires the fusion parameters such as the selection of the decision- thresholds corresponding to each modality and a fusion rule for their integration. The selection of these parameters depends upon the desired security requirement which can be expressed in terms of the two error rates: FAR and FRR. The choice of acceptable error rates is application dependent. For example, the high security applications (e.g., access to a building, or to a bank account) require the lowest possible

FAR (close to zero) but a permissible value of FRR to prevent false authentication (imposter) whereas the low security applications (e.g., public transport, classroom access) can be managed with somewhat high FAR but FRR must be stringent, i.e. as low as possible. The homeland security department envisages such a multi level security requirement in view of the forensic, civilian and high security applications within the purview of the color-coded terrorism threats [12].

However, most of the bimodal systems available in the literature [2]-[9] offer fixed level of security involving a fixed number of fusion parameters for the desired level of accuracy. The security provided by these biometric authentication systems is not adaptive to cater to the wide ranging applications requiring multiple levels of security [13]. Therefore, the need arises for an adaptive bimodal system capable of selecting the optimal fusion parameters to achieve the desired level of security. Hence an optimization approach is needed to guide the search for the optimum parameters. Some of the approaches for searching the optimal parameters can be categorized as: Deterministic methods [14], Probabilistic methods [15], and Evolutionary methods [16]. In comparison to other search algorithms, Evolutionary Computation (EC) techniques [16] generate a set of relevant solutions, termed as *population* and then find an optimal solution by searching and updating the memories (i.e. the past history of the particles) of the population. The techniques, such as Genetic Algorithm (GA) [16], Particle Swarm Optimization (PSO) [17], Ant Colony Optimization (ACO) [18], Bacteria Foraging (BF) [16] are generally influenced by the social behavior of species such as birds, fish, ants, bacteria, etc.

The advantages of the evolutionary computation include that it is easy-to follow, robust to changing circumstances, and flexible. Further, most of the evolutionary techniques provide a global solution which may not be the case with the optimization techniques which give the local optima. Evolutionary techniques can be applied to those real world problems for which the heuristic algorithms can lead to unsatisfactory results. Thus, the use of the evolutionary techniques is gaining immense popularity, particularly for the real life problems. This paper presents a novel bimodal system having the hand-veins and the hand-shapes as its two biometric modalities and employs ACO for selecting the fusion parameters of this system for meeting varying security levels.

The paper is organized as follows: a few state of the art methods on the bimodal/ multimodal systems are presented in Section II. The image acquisition and the normalization of both the modalities are briefed in Section III while the feature extraction is described in Section IV. The fusion methodology of the bimodal system using ACO is discussed in Section V. The experimental results and a relevant discussion are relegated to Section VI and VII respectively. Finally, the conclusions are given in Section VIII.

II. THE STATE OF THE ART

A. The Prior Works

The fusion of multiple biometrics in a multimodal system can be broadly categorized into two stages: (1) *fusion prior to matching*, which includes *sensor level* and *feature level* fusion schemes and (2) *fusion after the matching*, which can be performed at both the *score level*, and *decision level*. Though the extracted biometric features have rich information on the biometric data, the fusion schemes cause the deterioration of performance due to the loss of information at each level of fusion [1] [11]. As compared to the fusion prior to matching, the fusion after matching is not entangled with the issue of incompatibility. Hence, either the score level or the decision level

fusion scheme is preferred over the sensor level or the feature level fusion scheme in the multimodal biometric systems [11].

The history of bimodal/multimodal systems is more than a decade long and various developments have taken place during this period. Hong et al .[19] have proposed a bimodal system using fingerprint and face in 1998. They have fused the Eigen faces and the minutiae features from fingerprints at the decision-level to achieve FRR of 6.6% and FAR of 0.01% in comparison to FRR of 61.2% for face and 10.6% for fingerprint at the same FAR. Frischholz et al. [20] have presented probably the first commercial system BioID using lip movement and face images extracted from the same video sequence. They have achieved equal error rate (EER) of below 1% by their system. In the development of a multimodal biometric system, Fierrez et al. [21] have fused fingerprint, face, and signature. Using the three fusion schemes, viz., sum rule, user-independent SVM and user-dependent SVM, they achieve EERs of 0.5%, 0.3%, and 0.005% respectively. But without fusion the individual EERs are rather very high, i.e. 3%, 4% and 10% on fingerprint, face and signature respectively.

Kumar et al. [22] have come up with a bimodal system consisting of two biometric modalities, palmprint and hand-geometry. They have undertaken two types of fusion one at the feature level and other at the score-level. Their bimodal system yields FAR of 0% at FRR of 1.41% with the score-level fusion. Ross et al. [23] have attempted at a multimodal system using fingerprint, hand-geometry, and face as its constituent modalities. They have compared three fusion approaches: sum rule, decision tree, and linear discriminant function on their multimodal system. The sum rule is shown to give FAR of 0.03% at FRR of 1.78% while the fingerprint alone leads to FAR of 0.01% at FRR of 25%. Jain et al [24] have proposed another multimodal system using fingerprint, hand-geometry, and face biometrics. They achieve FAR of 0.1% at FRR of 1.4% on the multimodal system whereas fingerprint is alone found to offer FAR of 0.1% at FRR of 6.4%.

From the above we conclude that the bimodal/multimodal systems are surpassing over the unimodal systems in terms error rates. However, most of them offer a fixed level of security with a fixed number of fusion parameters. On the other hand, a biometrics-based multimodal system should be adaptive to cater to the wide ranging applications requiring multiple levels of security [13]. An adaptive bimodal/multimodal [25-26] biometric system is relatively new advancement and is of particular interest to the present work. Kalyan et al. [25] have presented an adaptive multimodal biometric system where they consider 16 binary decision rules arising from the two participating sensors and adaptively select one rule using binary PSO. In another approach for an adaptive multimodal system, Kumar et al. [26] have adopted the score-level fusion with four fusion rules; two of which are linear (sum and product) and the rest two are non-linear (exponential-sum and Tanh). They have implemented these rules on three bimodal systems consisting of i) palmprint and iris; ii) face and speech; and iii) fingerprint and hand-geometry. The performance of the score level fusion with their rules is shown to be better than that reported in [20]. However, both [25] and [26] utilize PSO to learn the fusion parameters.

B. The Motivations

The survey on bimodal/multimodal authentication systems [2]-[9] shows that most of bimodal/multimodal systems do not meet the varying security requirements [13]. Further, the hand-based biometrics such as fingerprint and palmprint, though possess significant details yet they are prone to imposter attacks as people leave their hand marks wherever they touch. These two multimodal systems in the literature [25-26], which offer an adaptive security, use PSO for

the optimal selection of the fusion parameters. The PSO algorithm has its own limitations like ending up in the local minimum leading to the premature convergence [27] and is applicable to the continuous domain. Further, the sigmoid function used to convert the continuous PSO into the binary PSO in [25][26] is a discrete variant of PSO which approximates the velocity and the position updates of each particle. Going in for a direct discrete domain evolutionary technique is certainly advantageous in place of the discrete variants of PSO. Thus there is a need to explore other evolutionary techniques for adaptive security. In this context, Bacterial foraging [18] that has a mechanism to decide the direction for the optimal solution may seem useful. However, the foraging strategies involve a large computation in four stages: *chemotaxis*, *swarming*, *reproduction*, and *elimination and dispersal* to reach upon an optimal solution. Hence they are not well suited for online biometric applications where quick decisions are of primary concern. Similarly genetic algorithm [28] has several misgivings like an expensive fitness function, convergence to local minimum, and prospectus of divergent solutions on dynamic datasets. We therefore, use ACO for the selection of fusion parameters in the proposed bimodal system.

C. *The Proposed Approach*

Despite the current popularity of the hand-vein patterns as a promising biometric trait [10][29]-[31], they are rarely used in bimodal/multimodal systems. This motivates us to develop an adaptive bimodal system incorporating the hand-vein patterns and hand-shapes that possess the contour information on the dorsal part of the hand. An attempt is already made in the creation of a bimodal system comprising hand-veins and the palmprints [32] based on the decision-level fusion facilitated by ACO. However, our preference to hand-shapes over palmprints is due to its touchlessness and less chances of imposter attacks. The block diagram of the bimodal authentication system is shown in Fig. 1.

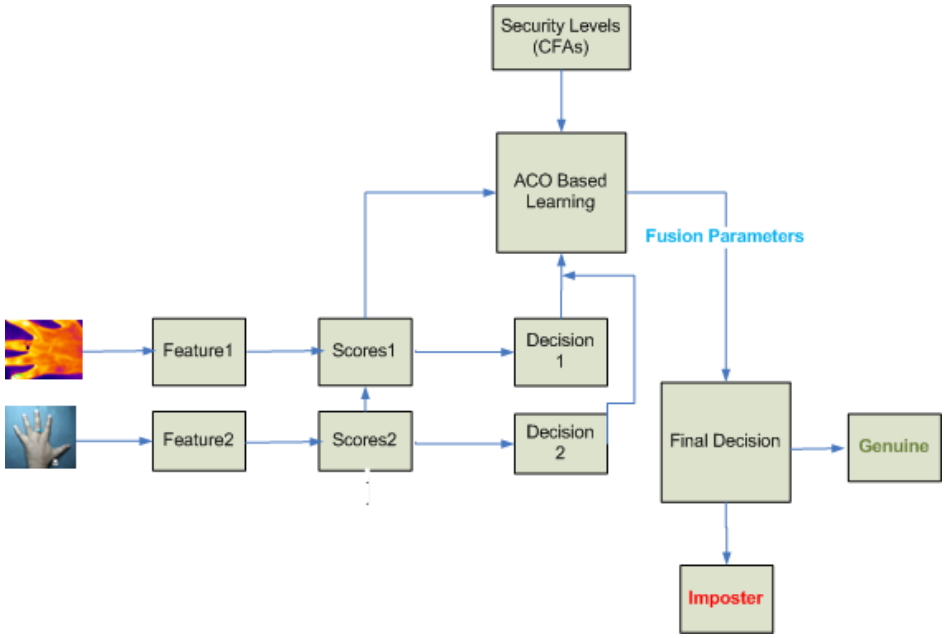


Fig. 1 Block diagram of the ACO-based adaptive bimodal system

The three-fold contributions of our work are now elaborated here:

1. Proposition of a bimodal system by fusing hand-vein patterns with the hand-shape at the score-level. The infrared thermal camera is used for the acquisition of vein patterns and

an ordinary digital camera for hand-shape. The Radon transform the linear trends in an image by aggregating the intensities along all possible lines [33]. The transform is used here for extracting features from the hand-vein patterns. The Independent Component Analysis (ICA) based appearance features are extracted from the hand-shapes [34]. The claimed identity is authenticated by fusing the scores computed from these two traits.

2. Utilization of ACO for the bimodal system adaptive to various civilian and forensic applications [13]. ACO has been utilized in earlier approaches for fusion in multispectral palmprints [35], feature selection in face images [36], and our earlier work on fusion of hand-vein and palmprints [32]. Note that in ACO, after updating the pheromone levels, the conventional way is to choose new paths for the next iteration from the complete set of candidate solutions. However, instead of trying out all candidate solutions, the ones with low probability can be discarded. This narrows down the search space and improves the quality of paths chosen by ants. Hence, the probability of selecting the next solution is constrained to lie between the lower and the upper values determined from the current local and global solutions. In this work, a framework is developed using the local and global updates carried out in ACO. This framework is then used for the optimal selection of fusion parameters in the bimodal system devised using hand-vein and hand-shape.
3. A comparison of the two scenarios of adaptive multimodal fusion. We first compare the ACO-based selection of fusion parameters with the PSO based ones in our bimodal system. We find ACO-based bimodal system yields lower error rates than that of PSO. The score-level fusion methodology in ACO is also compared with the decision-level fusion approach and is found superior to it in the selection of the fusion parameters.

III. IMAGE ACQUISITION AND NORMALIZATION

A. Vein Images

The vein patterns located beneath the skin on the dorsal part of the hand are not visible in the ordinary light while hair, moles, warts, etc, on the skin hinder an accurate acquisition by an ordinary digital camera. Further colors and attributes of fat/thick skins reduce the visibility. The acquisition is done at the indoor temperature in the range of 10^0C- 30^0C and at the relative humidity of 25%-50% as per the weather conditions. The visibility of the thermal imaging is affected at the outdoor environment and in such conditions segmentation of the vein patterns from background is very difficult [30]. Even when the images are acquired indoor, the visibility is sometimes affected due to sudden change in the temperature like: when a user leaves the outdoor environment to enter the indoor environment for imaging. To counter the above situation, we wait for a few seconds for the operating conditions to be restored before starting the acquisition again. The imaging setup and the acquired image are shown in Fig. 2.

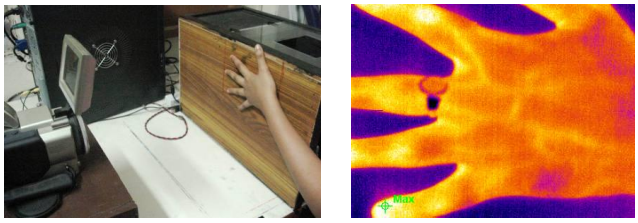
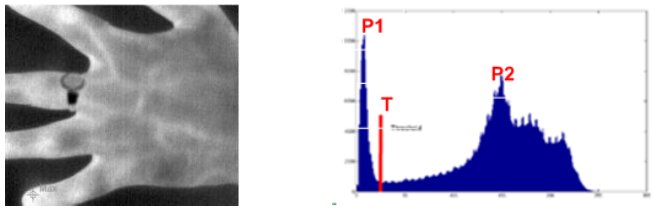


Fig.2. (a) Camera Setup (b) Captured Image

The images are acquired from the peg-free set up where the user's hand movement is not

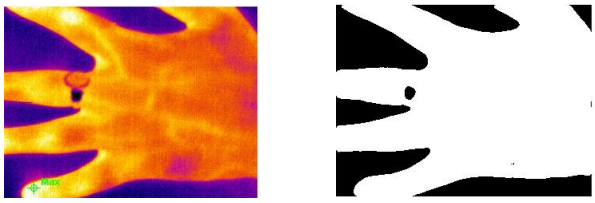
constrained by pegs and the ROIs are made invariant to rotation and translation. An algorithm is devised to extract Region of Interest (ROI) of vein patterns by fitting a coordinate system on the palm dorsa. The ROI extraction method, somewhat similar to [31], is as follows:

1. To determine the webbing between the fingers, the images are binarized. A histogram-based binarization method is utilized to choose the appropriate threshold from the captured images. By analyzing the bimodal property of the histogram of the dorsa images, local maxima P1 and P2 can be pinpointed and the local minimum T between them is determined. This local minimum becomes a threshold to segment the image. The input image and the corresponding histogram are shown in Fig. 3. Fig. 4 (a)-(b) shows the binarization using both the normal method and the histogram-based method.



(a) (b)

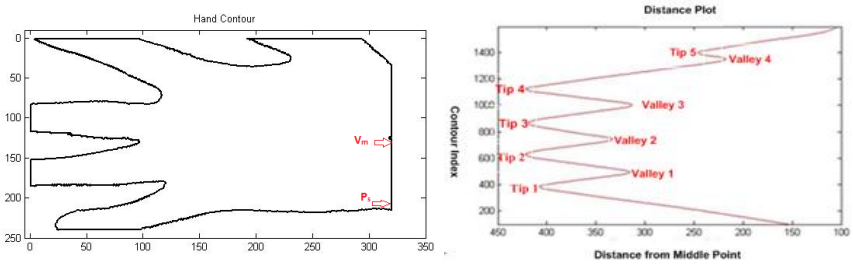
Fig. 3 (a) Vein image (b) Histogram of Image



(a) (b)

Fig. 4 (a) The acquired thermal image (b) The binarized image from histogram method

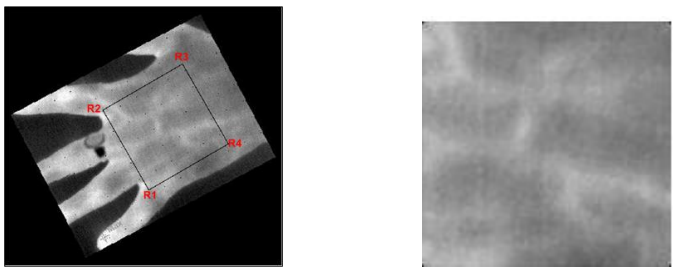
2. From the binarized image, a closed boundary is traced out in the clockwise direction. Moore-Neighbor tracing algorithm modified by Jacob's stopping criteria [44] is used for tracing the boundary points from the binarized image. The boundary points are then used for computing the tips and valleys of the hand by taking the bottom-left point marked P_s as the starting point and the middle point marked as W_m {Shown in Fig. 5 (a)}. We calculate the Euclidean distances between W_m and the boundary points P_s. The distances start increasing while tracing the boundary clockwise till the tips of the finger marked as point of local maxima are reached. These distances then start decreasing when calculated with the points way from the fingertips till the valleys of the same finger marked as local minima are reached. As evident from Fig. 5 (a) the distances are constant for a few boundary points (Represent the boundary of the fingertip) before start decreasing. We consider the midpoint of these points as tip of the finger. The valley point is the local minima. Similarly, five tips and four finger valleys are determined as shown in Fig. 5 (b).



(a) (b)

Fig. 5 (a) The boundary extracted from the hand contour (b) Tips and valleys of the hand

3. A tangent is drawn between the second and the fourth finger-valleys forming a line R1-R2. The middle point of this line is taken as the reference point and a perpendicular distance from this point decides the length of the square region {See in Fig. 6 (a)} with corners at R1, R2, R3 and R4. This square region is finally cropped and saved as ROI (See in Fig 6 b).



(a) (b)

Fig. 6 (a) Located ROI area on Palm dorsa (b) Extracted ROI

B. Hand Images

The images on the dorsal part of the hand are acquired using Canon A630 digital camera. A wooden box is made with an open cut on its upper part to fix the camera on it. The inner side of the box is meant to place a user's hand. The setup has a ring shaped fluorescent tube fixed under the roof of the box. The use of pegs and extra-illumination is avoided for ease of acquisition. The camera is focused on the dorsal part of the hand to capture the details. However, users are requested to place the hand in a marked place in order to keep it in the focus of the camera. The imaging setup and acquired image are shown in Fig. 7.



(a) (b)

Fig. 7 The Hand Image (a) The Camera Setup (b) The Captured Image

The tips and valleys from the acquired hand-shape help in the separation of the fingers from the hand. The separated fingers are rotated with respect to the pivotal axis formed by their vertices and the final hand contour is extracted by stitching them back to the wrist part [29]. The steps of the normalization are as follows:

1. The hand image is segmented from its background. As there is no extra light, the illumination management is applied; the acquired hand images are too difficult to be segmented by thresholding. Hence, a novel skin detection-based segmentation method is developed. In this, the red (R), green (G) , and blue (B) components of the image are made use of as under:

   Calculate

   $$Sum=R+G+B \quad (1)$$

   $$R_percent=R/Sum \quad (2)$$

   $$G_percent=G/Sum \quad (3)$$

   $$B_percent=B/Sum \quad (4)$$

   If

   $$\{ \text{R_percent, G_percent, B_percent} \} \in [30, 40]$$

   The pixel is on the skin

   Else

   The pixel is on the background

   After the skin detection, the image is binarized by making the skin surface white, while the background as black. The binarized image is then filtered to get the clean image. The original image and the binarized image are shown in Fig. 8 (a) and 8(b) respectively.



(a) (b) (c)

Fig. 8 Segmentation of the hand images: (a) The original image (b) The segmented by skin detection (c) The filtered image

2. The operations involving the separation of the fingers, rotation along the pivotal axis and stitching of the fingers to the wrist to get the normalized hand contour image are the same as in [34]. Fig. 9 shows these operations.

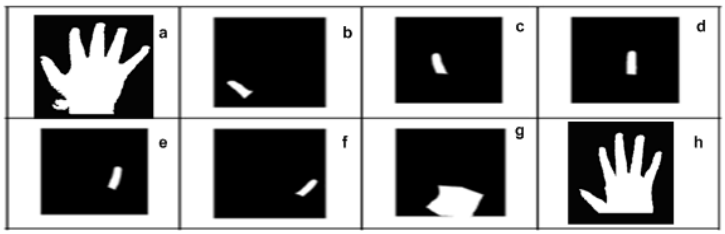


Fig. 9 Normalization of the Hand Image: (a) The binarized hand image (b)-(g) fingers and wrist separated from the hand image (h) The normalized hand-shape contour.

IV. FEATURE REPRESENTATIONS

A. Vein Features

The vein patterns possess enough texture details that serve as feature representation. One kind of texture features arise from the principal lines that can be used for matching two vein images. However, the vein patterns in our database are full of wrinkles which are small lines caused due to noisy acquisition; hence should be separated out from the principal lines before matching. The separation can be done by noting that the line energy on the principal lines is stronger than on wrinkles. Radon transform is used for detecting the directions of the principal lines and the energies in the vein patterns.

A straight line can be represented either in the slope-intercept form, y=ax + b in Cartesian co-ordinates, or by normal representation, $x\cos\theta+\sin\theta=\rho$ in the polar co-ordinates. Here ρ is the perpendicular distance from the origin and θ is the angle between the line and the *y*-axis. The Radon transform of a 2D function *f (x, y)* is a line integral along an arbitrary line in the xy-plane [37]. Copeland et al. [43] have presented a modified Radon transform to detect line segments in the local areas which are significantly shorter than the image dimensions by redefining the transform as:

$$Rad(r,\theta)=\int_{y\min}^{y\max}\int_{x\min}^{x\max} f(x,y)\times(x\cos\theta+y\sin\theta)dxdy \qquad (5)$$

The parameters: *x*max, *x*min, *y*max, and *y*min define a local area to perform Radon transform in (5). Radon based enhancement has also been implemented in [33] for extracting palmprint principal lines and in [37] for image representation. The approach of Matus et al. [37] utilizes the Radon transform for the finite length signals by the summing of the image pixels over a certain set of lines. The approach in [33] makes use of [37] for detecting the palm lines by modifying it using modulo operation to reduce the "wrap around" effect. Our implementation of Radon transform for vein patterns is similar to that in [37] for palmprints. They have modified the Radon transform for the real function *f* [*x*, *y*] by denoting a group $\{Z_p=0, 1. . . p-1\}$, where *p* is a prime number defined as:

$$Rad_f(k) = \frac{1}{c} \sum_{(i,j)\in L_k} f(i,j) \quad (6)$$

Where C is a scalar to control the scale of the transform, and L_k denotes a set of points constituting a line belonging to Z^2_p as:

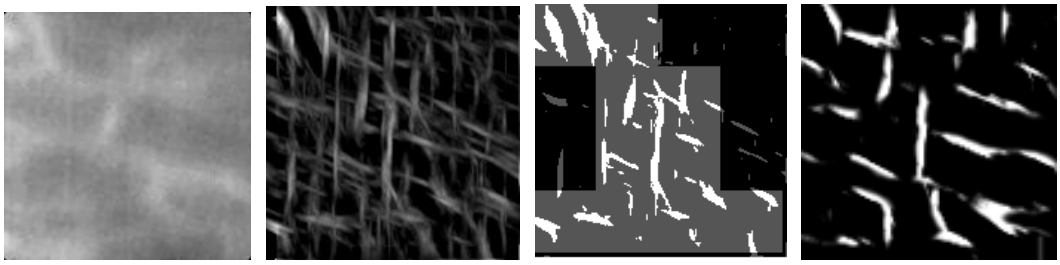
$$L_k = \left\{ (i,j) : j = k(i - i_0) + j_0 \middle| , i \varepsilon Z_p \right\}, 0 \le k \le p \quad (7)$$

Where (i_0, j_o) denotes the centre point of Z^2_p , and k indicates the slope of L_k. Here, Z^2_p can be thought of a local window in the vein ROI and L_k denotes a set of points extracted as principal lines from Z^2_p . We have another notation for L_k as $L(\theta_k)$ which represents the direction image and is calculated from:

$$\theta_{k(i_0, j_o)} = \arg(\min_k (Rad_f(k))) \quad (8)$$

$$\text{Energy}_{(i0,j0)} = (\min_k (Rad_f(k))) \quad (9)$$

For ROI of size $m \times n$ (Fig. 6 (b)), the centre of the window Z^2_p (p=17 in our case) moves from pixel to pixel. The energy image and the principal line image resulting from applying Eqn (9) and Eqn. (7)} [33] are shown in Fig. 10 (b)-(c). These images are combined to give an enhanced vein pattern image, shown in Fig. 8 (d) which is matched using the normalized Hamming distance by generating the matching scores.



(a) (b) (c) (d)

Fig. 10 (a) The Original ROI (b) Energy image (c) Principal line image (d) The Enhanced ROI

B. Hand-shape Features

The normalized hand-shapes (See Fig. 9) contain the contour information of the hand dorsal part. As the contour images have no textural or line details (as in the case of fingerprint or palmprint) it can be best represented by appearance features, which are obtained by the application of principle component analysis (PCA). However, as shown in [38] that the independent component analysis (ICA) is more robust and well suited as appearance features which have been applied to hand [34] and face [38] biometrics. In this work, the appearance features are extracted from hand-shape images by applying ICA which decomposes a mixed signal into the additive subcomponents of the non Gaussian type [39]. The dependencies in the hand-shape images need to be separated into independent components to obtain as ICA features. The theory and implementation of ICA is explained in [39]. The detail of ICA featuresis now in order.

The ICA model considers a hand-shape image as a combination of independent components (ICs) and a mixing matrix. Let $X_{K\times N}$ be the a hand-shape image, $M_{N\times N}$ be the mixing matrix and $C_{K\times N}$ be the matrix of ICs, these bear a relation:

$$X_{K\times N} = C_{K\times N} \times M_{N\times N} \quad (10)$$

The task in ICA is to estimate the independent components from the input image as:

$$C = X \times (M^{-1}) \quad (11)$$

Here, determining C from X is known as Blind Source Separation because the input source has to be separated into independent components without any knowledge of the mixing matrix. Though there are several approaches for the estimation of ICs from Eqn (10), we have considered the FastICA algorithm [39] that maximizes the statistical independence between the output components using maximization of the negentropy.

We have used here a hand-shape database of 150 users with 5 samples each. Each of the hand-shape images is of size 150×150. The mean-sample image for each user is computed by taking the average of 5 enrolled samples images. The mean-sample image is resized to get a image vector of size 22500×1. The mean-sample image is computed for all 150 users and the image matrix $X_{22500\times150}$ where K=22500 and N=150. As X is found to be of high dimensions, PCA is first employed on X for the dimensionality reduction. Let $U_{22500\times150}$ be the projection (eigenvector) matrix and $Mean_{22500\times1}$ is the mean vector arising from the application of PCA on X . The image matrix X is then transformed by PCA to get the projection matrix $Y_{150\times150}$ as follows:

$$Y_{150\times150} = (X_{22500\times150})^{T} \times U_{22500\times150} \quad (12)$$

By applying the FastICA algorithm on projected matrix Y we obtain the IC matrix $C_{150\times150}$ using Equations (10) and (11). At the verification/identification stages, let $Q_{22500\times1}$ be the image matrix computed from input hand-shape image. Subtracting the mean vector $Mean_{22500\times1}$ from it leads to

$$Vector1_{14400\times1} = Q_{22500\times1} - Mean_{225000\times1} \quad (13)$$

Multiplying the Vector1 with the PCA projection matrix gives the Eigen feature vector as:

$$Eigenfeature_{1\times150} = Vector1_{22500\times1}^{T} \times U_{22500\times150} \quad (14)$$

Use of the IC matrix on the Eigenfeature vector yields the ICA feature vector as:

$$ICAfeature_{150\times1} = B_{150\times150} \times Eigenfeature_{1\times150}^{T} \quad (15)$$

Here the Euclidean distance is used to match two ICA feature vectors. Let $ICAV_1$ and $ICAV_2$ be the two ICA features computed using Eq. (15). The matching score s between these two vectors is computed as:

$$s = \sqrt{\sum_{i=1}^{150} \{ICAV_1(i) - ICAV_2(i)\}^2} \quad (16)$$

V. ACO-BASED FUSION

A. Adaptive Bimodal management

In the proposed bimodal system, an optimal selection of fusion strategies along with the associated fusion parameters for meeting the multiple levels of the security requirements.

Decision-threshold and a fusion rule are the two basic *fusion parameters* that need to be selected optimally. The performance of a typical biometric system can be evaluated in terms of the error rates: FAR and FRR. One way of ensuring different security levels is to assign different costs to these error rates and study the fusion parameters for each security level. If high security is required, the cost of FAR can be chosen much higher than that of FRR; as in this case an imposter acceptance should be as less as possible. But, in the case of low security, the cost of FRR is higher than FAR; as in this case the genuine rejection is required to be as less as possible. Therefore, different security levels in a multimodal system can be enumerated by using the cost of false acceptance (CFA) and cost of false rejection (CFR). Both costs can belong to [0, 1] in the step size of 0.1 providing us with 20 discrete points which represent the different security levels. The fusion parameters need to be computed for each security level by employing error rates of the individual biometrics and the allowable fusion rules. The overall error (also our objective function) can be computed by taking both the costs as the weights to the FAR and FRR in the following:

$$G = CFA \times FAR + CFR \times FRR \tag{17}$$

Where $CFA + CFR = 2$

The FAR and FRR are the outcome of applying a decision-threshold *(α)* on the matching scores [26]. Four score-level fusion rules: Sum, Product, Exponential Sum and Tan-hyperbolic (Tanh) Sum are considered for fusion at the matching score-level as in [26]. Let s_j be the matching scores from j^{th} modality and W_j be the corresponding weight, where j=1, 2 for two modalities. The fusion rules for aggregating the matching scores are chosen as:

$$Sum = \sum_{j=1}^{2} s_j W_j \tag{18}$$

$$\text{Product} = \prod_{j=1}^{2} s_j^{W_j} \tag{19}$$

$$Exponential\ Sum = \sum_{j=1}^{2} \exp(s_j) W_j \tag{20}$$

$$Tanh\ Sum = \sum_{j=1}^{2} \tanh(s_j) W_j \tag{21}$$

Here, weights are selected according to the security levels for the adaptive fusion of the vein patterns and the hand-shape biometrics. The inputs to the ACO are: the security levels, the weights corresponding to both the modalities; and the four score-level fusion rules {Equations (18)-(21)}. The outputs of the ACO are: a decision-threshold, optimal weights corresponding to each modality, and a fusion rule. The task of ACO is to minimize the overall error *G* (Eq. 17) by selecting the weights for each modality and a fusion rule.

The population in ACO is initialized in the *D-dimensional* space. An ant can be represented as $X_{md} = (x_{m1}, x_{m2} \ldots x_{mD})$, where the first subscript *m* represents a m^{th} solution, and D denotes the dimension. For the score-level fusion, each ant has *"N+2"* dimensions, where *N* is the number of modalities which 2 in this work. Thus, out of four dimensions, the first 2 dimensions are meant for the weights (*W*) assigned to each modality, the 3^{rd} dimension is kept

for the decision-threshold, and 4^{th} dimension represents a fusion rule. An ant thus represents a parameter vector:

$$X_{md} = \{ W_1, W_2, \alpha, F\}, \qquad (22)$$

Where F is one fusion rule chosen from Eqns. (17)-(20).

B. A Brief Description of ACO

Ant colony optimization (ACO), first introduced by Marco Dorigo, is a class of optimization algorithms applicable to the problems seeking optimal paths [40]. The salient features of ACO include the positive response, circulated computation, and greedy heuristic. The positive response promotes quick search for optimal solutions, the circulated computation removes the premature convergence, and the greedy heuristic assists in searching for desirable solutions at the early stages of the algorithm. The ACO algorithm relies on the pheromone-based probabilities to search for the optimal paths [40]. While searching, the moving ants (probable solutions) deposit some pheromone which evaporates over a trail period on the chosen paths. Any ant encountering a previous trail decides to follow a path having a high concentration of pheromone meaning thereby many ants had already tread that path. As a result, the collective behavior emerging from several trails is a kind of positive feedback for the selection of an optimal path. The pheromone evaporation prevents the premature convergence as no single ant can ever dictate a path; and the selected path is a collective judgment [40].

The ants are initialized as in Eq. (22). Initially, each ant deposits an equal amount of pheromone level. Now, the objective function G {See Eq. (17)} is evaluated for each ant. The individual weights are applied oin the chosen fusion rule to get the fused matching score. The decision-threshold α is then used to compute FAR and FRR from the fused matching scores incorporating a security level. The updated pheromone level τ_i^{t+1} at $(t+1)^{th}$ iteration where i denotes the i^{th} ant is given by

$$\tau_i^{t+1} = \rho \times \tau_i^i + \frac{Q}{G}, \text{ if } i^{th} \text{ ant minimizing the } G$$
$$= \rho \times \tau_i^i, \text{ Otherwise} \qquad (23)$$

If any of the ants (newly computed solutions) participates in the minimization of the objective function (G), the value of the pheromone level increases by *Q/G*; where Q is a constant chosen by trial and error in the range [0.005, 0.01]. A value of 0.01 is found to be appropriate here. Along with the update of the pheromone levels, the probability of the i^{th} ant is computed for the next iterations as:

$$P_i = \frac{\tau_i^{(t)}}{\sum_{k=1}^{D} \tau_k^{(t)}} \qquad (24)$$

The probability of selecting ants which minimizes G gets better than that of other ants in Eq. (23) in the next iteration where the pheromone levels are comparatively less.

One important parameter of ACO is the evaporation factor (ρ) in Eq. (23) that influences the probability in Eq. (24). The ACO suffers from the problem of saturation of the pheromone

level for small values of ρ (close to 0) forcing the ants stuck on a particular path. On the other hand if ρ is close to 1, every path will have equal probability with no scope for optimization. To overcome this, ρ is set to 1 initially and subsequently decreased in steps of 0.005. This strategy not only prevents the ants from being dragged on any non-optimal path but also allows them to navigate to an optimal path. Unlike the traditional ACO where complete search space is used to select the solutions at the next iteration, the search space is constrained by introducing the lower and the upper values, i.e. *L* and *U*. The local best, the global best, and the current position (denoted by A_{lb} , A_{gb} and A_{cp} respectively) are calculated for all ants at every iteration. From A_{lb}, $A_{gb,}$ and A_{cp}, *L* and *U* are derived as the limiting values. At the first iteration, an equal amount of pheromone is assigned to all the paths and the probabilities are calculated accordingly. From the next iteration onwards, the probabilities are assigned by modifying (24) as :

$$p_i = \frac{\tau_i^{(t)}}{\sum_{k=L}^{U} \tau_k^{(t)}} \qquad (25)$$

The constraints on A_{cp}, A_{lb} and A_{gb} that facilitate the option of L and *U* are given in Table 1. The limiting conditions are checked for each ant and the values of L and U are decided accordingly.

Table1: The Limiting Conditions for the ACO search

Limiting conditions	The value of L	The value of U
$A_{lb} \in]A_{cp}\ A_{gb}[$	A_{cp}	1
$A_{cp} \in]A_{lb}\ A_{gb}[$	A_{lb}	1
$A_{gb} \in]A_{lb}\ A_{cp}[$	A_{lb}	A_{cp}
$A_{gb} \in]A_{cp}\ A_{lb}[$	A_{cp}	A_{lb}
$A_{cp} \in]A_{gb}\ A_{lb}[$	0.001	A_{lb}
$A_{lb} \in]A_{gb}\ A_{cp}[$	0.001	A_{cp}

C. The ACO Algorithm

The ACO algorithm is outlined here:

I. Initialize the population of the ants randomly in the search space (See Section *V A*). For the score-level fusion, the weights corresponding to each modality are assigned values between 0 and 1.

II. Specify the pheromone levels of all ants as the discrete array $(\tau W_1 , \tau W_2 ,.., \tau W_P)$ for the score-level fusion. They are fixed at the same pheromone value at the first iteration and the probabilities are computed using Eq. (24).

III. Update the local and global best positions of the ants as in Section *III B* and evaluate the probabilities using Eq. (25).

IV. Iterate until the end of the specified number of iterations.

VI. EXPERIMENTAL RESULTS

A database of 150 users with 5 samples per user is acquired at the Biometric Research Laboratory, IIT Delhi. The Radon features for the vein patterns and ICA features for Hand-shapes are extracted for all the enrolled samples of each user. Out of 5 enrolled images, 3 samples are chosen for the training and the rest are for the testing. The Euclidean distance is used to match both the Radon and ICA features extracted from vein patterns and hand-shape respectively (See Eq. 15). Some of the conventional performance metrics and the key terms utilized in the proposed bimodal system are elaborated as follows:

I. *The matching scores*: The training and the testing samples of each user are matched using Euclidean distance to calculate the matching scores. If the training and the testing images belong to the same user, the matching scores are referred to as the *genuine score*, otherwise *imposter scores*. Hence, in case of genuine matching score, all the three training images of a user are matched with first test sample of the same user to compute three matching scores. The minimum of these scores is taken as genuine score from first test sample. Similarly a matching score for another test image is computed for the same user. Hence, for 150 users, 300 genuine matching scores (150×2) are computed for both vein patterns and hand-shapes. For imposter matching score, all the three train images of a user are matched with the (149×2) test images of another 149 users. The minimum of out of the three matches is stored as imposter score corresponding to each of the test user. Hence, for 150 users, 44700 (150×149×2) are computed for both the vein patterns and hand-shapes.

II. T*he error rates:* These result from the comparison of both genuine and imposter scores against a threshold. The percentage of the rejected genuine matching scores by the decision-threshold is termed as false rejection rate (FRR) while the percentage of accepted imposter scores is termed as false acceptance rate (FAR). For a typical biometric authentication, the minimum FAR must be specified to get the corresponding genuine acceptance rate (GAR) =100-FRR.

III. *The Receiver Operating Characteristics (ROC):* It is a plot of FAR Vs. GAR obtained by varying the threshold and finding FARs and GARs. In the conventional biometric systems, the threshold that yields the lowest FAR is selected from ROC. This threshold is used for making the decision of acceptance and rejection during the verification stage. We have used this curve to show the performance of the hand shape and the hand veins as individual biometric modalities. However, instead of ROC, our bimodal system employing the two traits uses ACO for the optimal selection of fusion parameters and decision threshold for each security level.

IV. In our work, the selection of the optimal fusion parameters is largely dependent on the cost factor CFA; so the algorithm is run 100 times as in [25] for each value of CFA. The parameters that minimize G in Eq. (17) are varied along with CFA. However CFA is varied in the interval [01, 1.9]. The average and the standard deviation of the minimum error over 100 runs (which is sufficient for all practical purposes) are needed since the average minimum error helps ascertain the performance of the bimodal system [25] [26] whereas the standard deviation of the minimum error helps to know the consistency of the algorithm [26].

The results of the bimodal system are discussed in the following order: First, the results of the score-level fusion using ACO; Second, comparison of ACO with PSO; and Third, comparison of the score-level fusion with the decision-level fusion.

A. The Score-level Fusion using ACO

The histograms of matching scores of the hand-shapes and hand-veins are shown in Fig. 11 (a)-(b) respectively. The ROCs of these modalities appear in Fig. 11 (c). The plots of the average and the standard deviation of the minimum error over 100 runs of the algorithm are shown in Fig. 12 (a) and (b) respectively. The average error of our bimodal system is significantly low as compared to that of the separate unimodal systems {Please see ROC curves in Fig. 11 (c)}. The standard deviation of the error is found as 0.0007 which shows that the ACO algorithm is stable over the 100 runs of algorithm. In order to see the suitability of the fusion rules, the probabilities of selection of each rule over 100 runs of the algorithm for each value of CFA are displayed in Fig. 13. From this, we can find that most of the times SUM rule and the PRODUCT rule are selected as the best fusion rule by the ants for most of the weights. However, SUM rule has shown slightly better performance over PRODUCT rule. As evident from the figure, PODUCT rule performs better than the SUM rule when CFA less than 0.5, though, at 0.3 the SUM rule performs better. At 0.6 both the rules have almost the same probability of selection. When CFA increases from 0.6, the SUM rule is selected with the highest probability implying high performance for the rest of the values of CFA.

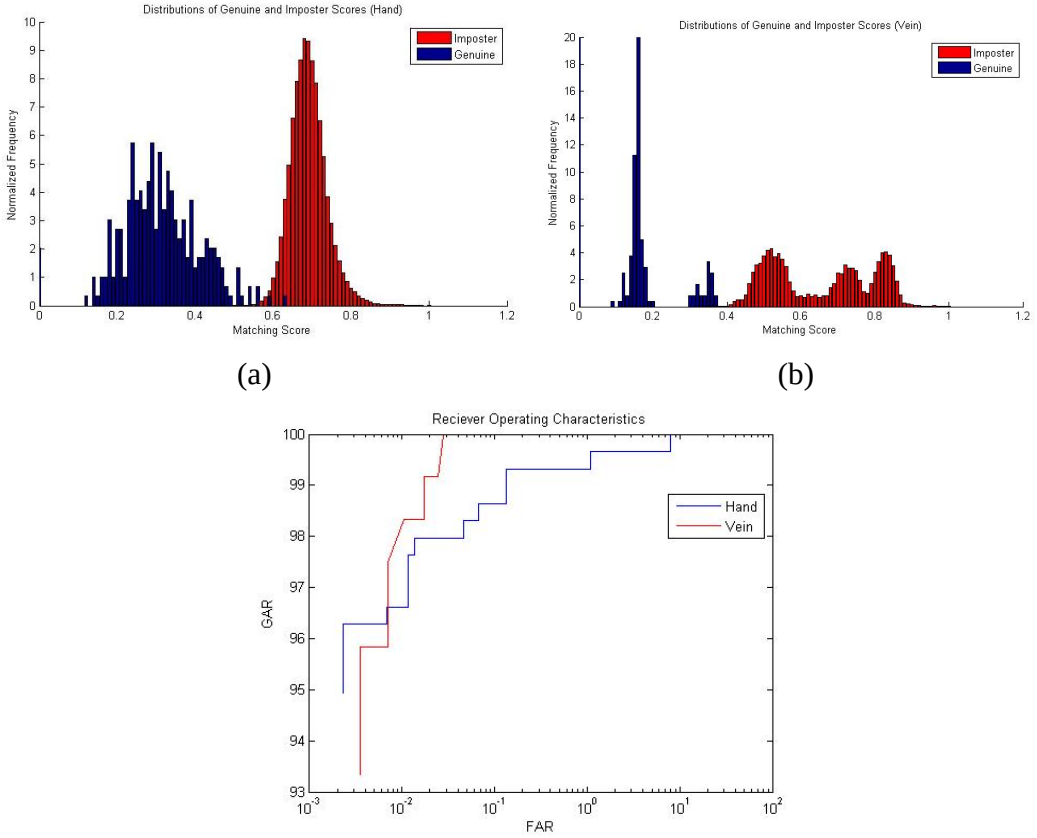
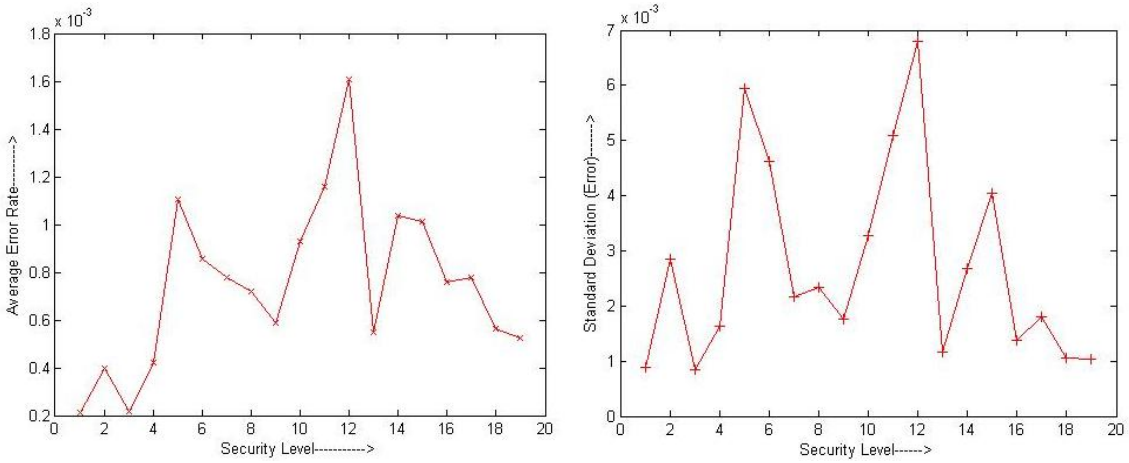


Fig. 11: (a) Histogram of the matching scores of the hand-shape (b) Histogram of the matching scores of vein patterns (c) Combined ROC curve for both



(a) (b)

Fig.12: The plots of (a) the average of error (b) the standard deviation of error

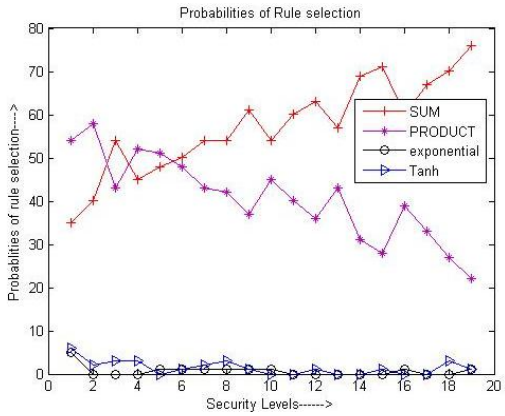
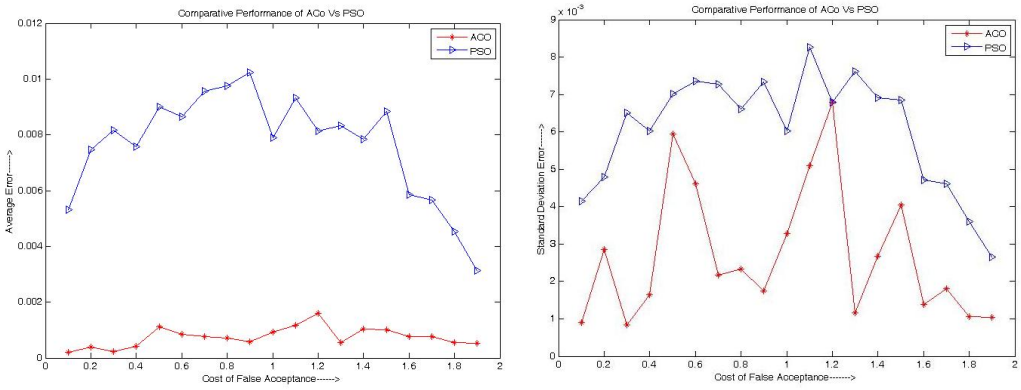


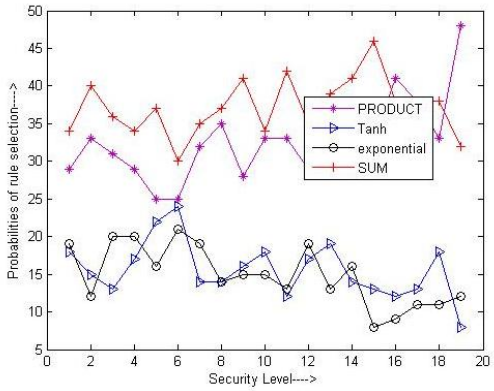
Fig. 13 Probabilities of the rule selection for each value of CFA

B. Comparison of ACO with PSO

PSO was used for the selection of fusion parameters in the previous works on the adaptive multimodal fusion [25]-[26]. In PSO, the parameters used are: inertia weight (W) = 1.2, the social (ψ_1) and the cognitive (ψ_2) parameters are 1.2 and 0.9 respectively. The number of swarms is 15 whose initialization is the same as that of ants (See Section *V-A*) and 150 iterations are provided for getting a converged solution using PSO algorithm. The ACO Vs PSO plots of the average of error and the standard deviation of error are shown in Fig. 14 (a) and Fig. 14 (b) respectively for the score-level fusion. In both the plots, ACO-based selection of fusion parameters is found superior to that of PSO. The plot of the probabilities of rule selection for each CFA is shown in Fig. 14 (c). Unlike ACO, the probabilities of the selection of exponential and Tanh rules are higher with PSO.



(a) (b)

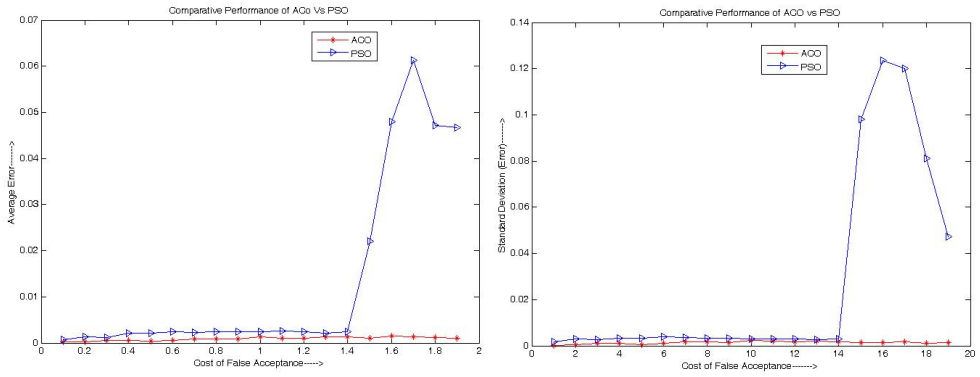


(c)

Fig.14: The plot of PSO Vs. ACO for the score-level fusion (a) The Average (b) The SD; (c) Rule selection using PSO

C. Comparison of Score-level fusion with Decision-Level Fusion

The decision-level fusion strategy is employed in the PSO-based multimodal system first in [25] and then in [26]. However, the results of [26] indicate that the decision-level fusion scheme is inferior to the score-level fusion if the selection of fusion parameters is by PSO. Our study on the bimodal system reveals two points: (1) The decision-level fusion with ACO-based parameter selection is superior to the PSO-based selection, and (2) The score-level fusion is better than the decision-level fusion with ACO. As a verification of these points, see the plots of ACO Vs. PSO for the average error and the standard deviation of the error in Figs. 14(a)-(b) and Figs. 15 (a)-(b) respectively. To justify the first point, we note that ACO incurs less error in both the plots. Moreover, the differences of the average errors and of the SD errors between the ACO and PSO due to the score-level fusion (See Figs. 14a-b) are more than those due to the decision-level fusion {Fig. 15 (a)-(b)}. As regards the second point, see the average error for each value of CFA in Fig. 16 (a) and SD error for the same value of CFA in Fig. 16 (b). In both of the plots, score-level fusion gives low error rates than the decision-level.



(a) (b)

Fig.15 The plot of PSO Vs ACO for the decision-level fusion: (a) The Average (b) The SD

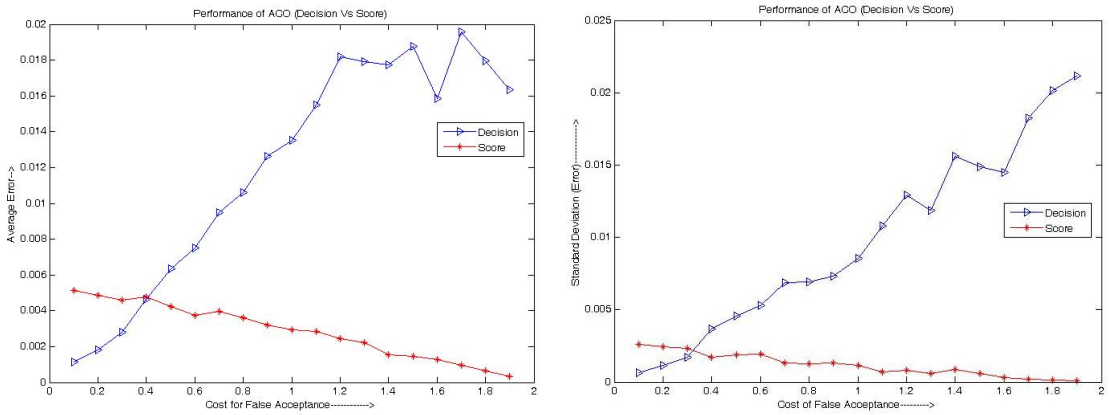


Fig. 16 The Performance of ACO on the Decision Vs. Score-level fusion (a) The Average (b) The SD

VII. DISCUSSION

A survey on the biometric authentication shows diverse performance of biometric modalities based on application, availability, and user convenience. The eye- based biometrics like, iris or retina possesses high details and individuality. However, being a small and small target, irises are difficult to acquire with low user convenience. The face biometric has high user convenience and availability but it suffers from large pose and expression variations. The advantage with hand-based biometrics is that it possesses high textural details with high user continence and availability. Though, people generally leave their handprints unconsciously by touching and chances of imposter attacks are much more in such scenario. Therefore, the recent trend in biometric literature is to investigate biometric modalities from the dorsal part of the hand which is acquired touch-less having lesser chances of imposter attacks.

Knuckle print is one such potential biometric modality recently developed with promising recognition rates. Kumar et al. [41] have achieved near 100% accuracy with FAR of 0.03%. Our earlier work [42] on knuckle print verification makes use of the fuzzy binary tree and ACO. Vein

patterns are another biometric modality that caught the attention of several researchers [10][30]-[31]. Hidden beneath the skin, vein patterns have lesser chances of imposter attacks and are shown to yield high recognition rates. Wang et al. [30] have shown 0% FAR and 0% FRR on infrared vein patterns. However, their experiments are conducted on a small database of only 30 users with 9 samples each. moreover, infrared thermal vein patterns need somewhat controlled environment at the time of acquisition along with proper management of imaging [30]. But once acquired correctly, they possess high textural details.

Our earlier works on unimodal [31] and bimodal [32] systems using hand vein patterns are encouraging. Though the hand vein patterns are less investigated, the bimodal system developed by integrating the vein patterns and hand-shapes is new. This work targets the indoor security applications and may be well suited for situations like *the access authentication for secret labs/building*. It is well known that when the user identity is known, verification is the most preferred mode for the personal authentication and some contributions in this mode are [8]-[10][13][25][26]. The case of identification is more suitable to the forensic and surveillance applications. However, in most of the civilian applications like building access problem or employee attendance system, a user can have his identity with him. The task of authentication in such cases is to verify the request of a claimed identity. A typical biometric system has therefore been designed by learning a suitable decision threshold which can reject an imposter attempt and at the same time accept a genuine attempt. The selection of an optimal decision threshold depends upon the expected security requirements which can be expressed in terms of the two error rates: false acceptance rate (FAR) or false rejection rate (FRR). The choice of acceptable error rates is application dependent (See Section *I-B*). The motivation for this work stems from the desire to design a bimodal system using infrared thermal hand veins and had shape by optimally selecting the decision parameters using ACO. It is noted that different security levels for different applications can be achieved using CFA/FFR (See Section *V-A*) and for each security level an optimal decision threshold, fusion rule, and weights for score level fusion {Eqn (18)-(21)} are computed using ACO (See Section *V-B*) in the proposed bimodal system. The computed decision thresholds can be utilized to distinguish a genuine attempt from an imposter attempt.

VIII. CONCLUSIONS

This paper presents a novel bimodal system using infrared thermal hand-veins and visible hand-shape modalities. ACO is used to fuse these modalities at the score-level by learning the fusion parameters. Both traits can be simultaneously extracted from the user's hand, and the claimed identity is authenticated by integrating them. In order to make the proposed system adaptive to the varying security levels, ACO is used to select the fusion parameters according to the user defined cost functions. The experimental results indicate that for the varying security levels, the ACO-based selection of fusion parameters is well suited to the bimodal system as it fares well over PSO.

The experimental results carried out on the IITD database of 150 users demonstrate that the fusion parameters selected using ACO yield the lower values of the average error and the standard deviation error than those of the PSO. The main reason for the superior performance of ACO is that it avoids the premature convergence which in turn leads to near optimal selection of the fusion parameters most of the times. Further, ACO is defined for discrete domain and

therefore may be a better option than discrete variants of PSO like, binary PSO used in earlier systems.

The performance of the proposed bimodal system has the potential to authenticate reliably because of the internal structure of the vein patterns, that possess the properties such as uniqueness, stability and strong immunity against forgery, and the external structure of the hand-shape, which provide an effective appearance features. Thus our bimodal system is unique in the sense of capturing both internal and external structural information. As far as the implementation is concerned, our authentication system works on the computer with the configuration Core 2 Quad, 4 GB RAM using MATLAB 2009.

REFERENCES

[1] K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[2] A.K. Jain, L. Hong and R. Bolle, "On-line Fingerprint Verification", *IEEE Transactions on PAMI*, Vol. 19, No. 4, pp. 302-314, 1997.

[3] L. Hong, Y. Wan and A.K. Jain, "Fingerprint Image Enhancement: Algorithms and Performance Evaluation", *IEEE Transactions on PAMI* , Vol. 20, No. 8, pp.777-789, August 1998.

[4] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the Individuality of Fingerprints", *IEEE Transactions on PAMI*, Vol. 24, No. 8, pp. 1010-1025, 2002.

[5] D. Zhang, W. K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, Sep. 2003.

[6] J. You, W. Li, and D. Zhang, "Hierarchical palmprint identification via multiple feature extraction," *Pattern Recognition*, vol. 35, pp. 847-859, 2002.

[7] A. K. Jain and J. Feng, "Latent Palmprint Matching", *IEEE Trans. PAMI* vol. 31, no. 6, pp. 1032-1047, June, 2009.

[8] A. K. Jain and N. Duta, "Deformable matching of hand shapes for verification", *Proc. IEEE International Conference on Image Prcoessing*, October 25-28, Kobe, Japan, 1999.

[9] A.K. Jain, A. Ross and S. Pankanti,"A Prototype Hand Geometry-based Verification System", *2nd Int'l Conference on Audio- and Video-based Biometric Person Authentication (AVBPA)*, Washington D.C., pp. 166-171, March 22-24, 1999.

[10] Chih-Lung Lin, and Kuo-Chin Fan, "Biometric verification using thermal images of palm-dorsa vein patterns", *IEEE transactions on circuits and systems for video technology*, Vol. 14, pp. 199-213, 2004.

[11] K. Nandakumar, "Integration of multiple cues in biometric systems", *PhD thesis*, Department of Computer Science & Engineering, Michigan State University, East Lansing, USA, 2005.

[12] Available at: http://www.dhs.gov/xlibrary/assets/CitizenGuidanceHSAS2.pdf

[13] A. Kumar, "Dynamic security management in multibiometrics," in Multibiometrics for Human Identification , B. Bhanu and V. Govin-daraju, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[14] R. Horset and H. Tuy, "Global Optimization – Deterministic Approaches", *Springer*, NewYork, 1996.

[15] Noga Alon and Joel Spencer, "The Probabilistic Method", *John Wiley*, June 2008 .

[16] T. Bäck, D. Fogel and Z. Michalewicz, "Handbook of Evolutionary Computation", IOP Publishing and Oxford University Press, New York, 1997.

[17] J. Kennedy and R. C. Eberhart, "PSO optimization," *Proc. IEEE Int. Conf. Neural Networks*, pp.1941–1948, 1995.

[18] Kevin M Passino, " Biomimicry of Bacteria Foraging for Distributed Optimization and control", IEEE *Control Systems Magazine*, June 2002.

[19] L. Hong and A. K. Jain, "Integrating faces and fingerprints for personal identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 12, pp. 1295–1307, Dec. 1998.

[20] R. Frischholz and U. Dieckmann, "BiolD: A multimodal biometric identification system," *Computer*, vol. 33, no. 2, pp. 64–68, Feb. 2000.

[21] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero, and J. Gonzalez-Rodriguez, "A comparative evaluation of fusion strategies for multimodal biometric verification," in *Proc. 4th Int. Conf. Audio- Video-Based Biometric Person Authentication*, J. Kittler and M. Nixon, Eds., 2003, vol. LNCS 2688, pp. 830–837.

[22] A. Kumar, D. C. M. Wong, H. C.Shen1, and A. K. Jain, "Personal verification using palmprint and hand geometry biometric,"in *Proc. 4th Int.Conf. Audio- Video-Based Biometric Person Authentication*, J. Kittler and M. Nixon, Eds., 2003, vol. LNCS 2668, pp. 668–678.

[23] Arun Ross, Anil Jain, "Information Fusion in Biometrics", *Pattern Recognition Letters*, vol. 24 pp. 2115–2125, 2003.

[24] A. K. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, 2005.

[25] Kalyan Veeramachaneni, Lisa Ann Osadciw, and Pramod K. Varshney, "An Adaptive Multimodal Biometric Management Algorithm", *IEEE Trans. On Systems, Man, and Cybernetics—Part C: Applications and Reviews*, vol. 35, no. 3, August 2005.

[26] Ajay Kumar, Vivek Kanhangad, and David Zhang, "A New Framework for Adaptive Multimodal Biometrics Management", *IEEE Trans. on Information Forensics and Security*, Vol. 5, no. 1, pp. 92-102, March 2010.

[27] E.S. Peer, F. mi de11 Bergll, A.P. Engelbrecht, "Using neighbourhoods with the guaranteed convergence PSO", *IEEE Proceedings of Swarm Intelligence*, pp. 235-242, USA, April, 2003.

[28] Genetic Algorithm, Available at: http://en.wikipedia.org/wiki/Genetic_algorithm

[29] Y. Ding, D. Zhuang and K. Wang, "A study of hand vein recognition method," *Proc. IEEE Intl. Conf. Mechatronics & Automation*, Niagara Falls, Canada, pp. 2106 – 2110, Jul. 2005.

[30] L. Wang, G. Leedham, and S.-Y. Cho, "Infrared imaging of hand vein patterns for biometric purposes", *IET Compt. Vis.*, Vol.1, pp. 113-122, 2007

[31] Amioy Kumar, M. Hanmandlu and H. M. Gupta, "Online Biometric Authentication Using Hand Vein Patterns", *IEEE Symposium: Computational Intelligence for Security and Defense Applications*, Ottawa, Canada, 8-10 July 2009.

[32] Amioy Kumar, Madasu Hanmandlu, Harsh Sanghvi, H. M. Gupta, "Decision level biometric fusion using Ant Colony Optimization", *In Proceedings of ICIP,* pp.3105-3108, HongKong, Sept. 2010.

[33] De-Shuang Huang , Wei Jia, David Zhang, " Palmprint verification based on principal lines", Pattern Recognition, Vol. 41, pp. 1316 – 1328, 2008.

[34] Erdem Yoruk, Ender Konukoglu, and Bulent Sankur, "Shape-Based Hand Recognition", *IEEE Transactions On Image Processing*, Vol. 15, No. 7, JULY 2006.

[35] D. R. Kisku, P. Gupta, J. K. Sing, and C. J. Hwang, "Multispectral Palm Image Fusion for Biometric Authentication using Ant Colony Optimization", *IEEE 1st International Workshop on Emerging Techniques and Challenges for Hand-based Biometrics (ETCHB 2010)*, pp. 1-7, Istanbul, Turkey, August 2010.

[36] H. R. Kanan, K. Faez, M. Hosseinzadeh, "Face Recognition System Using Ant Colony Optimization-Based Selected Features", *IEEE Symposium on CISDA*, PP. 57-62, 1-5 April 2007.

[37] F. Matus, J. Flusser, "Image representations via a finite Radon transform", IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 996–1006.

[38] Marian Stewart Bartlett, Javier R. Movellan ,and Terrence J. Sejnowski, "Face Recognition by Independent Component Analysis", *IEEE Transactions on Neural Networks*, Vol. 13, No. 6, November 2002.

[39] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Netw.*, vol. 13, no. 4–5, pp. 411–430, 2000.

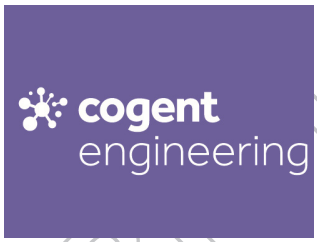
[40] Dorigo M., G. Di Caro and L. M. Gambardella "*Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem*" IEEE transactions on evolutionary computation, vol. 1, no. 1, April 1997.

[41] Ajay Kumar, Yingbo Zhou, "Personal Identification using Finger Knuckle Orientation Features", *Electronics Letters*, vol. 45, no. 20, September 2009.

[42] A. Kumar, M. Hanmandlu, and H. M. Gupta, "Ant colony optimization based fuzzy binary decision tree for bimodal hand knuckle verification system", *Expert Systems with Applications*, 2012, HUhttp://dx.doi.org/10.1016/j.eswa.2012.07.042UH. {In the Press}.

[43] A.C. Copeland, G. Ravichandran, M.M. Trivedi, Localized radon transform-based detection of ship wakes in SAR images, IEEE Trans. Geosci. Remote Sensing 33 (1) (1995) 35–45.

[44] T. Pavlidis, *Algorithms for Graphics and Image Processing*, Computer Science Press, Rockville, Maryland, 1982.



A novel fuzzy energy based level set method for medical image segmentation

Mahipal Singh Choudhry, Rajiv Kapoor

Accepted Manuscript Version

This is the unedited version of the article as it appeared upon acceptance by the journal. A final edited version of the article in the journal format will be made available soon.

As a service to authors and researchers we publish this version of the accepted manuscript (AM) as soon as possible after acceptance. Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). Please note that during production and pre-press, errors may be discovered which could affect the content.

© 2018 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

Publisher: Cogent OA

Journal: *Cogent Engineering*

DOI: http://dx.doi.org/10.1080/23311916.2018.1475032

ACCEPTED MANUSCRIPT

A Novel Fuzzy Energy Based Level Set Method for Medical Image Segmentation

1. First Author and Corresponding Author: Mahipal Singh Choudhry, Associate Professor, Department of Electronics & Communication Engineering, Delhi Technological University, Delhi, India.
   E-mail: msc_1976@yahoo.com
2. Second Author: Rajiv Kapoor, Principal, Ambedkar Institute of Advanced Communication Technologies & Research, Delhi, India.
   E-mail: rajivkapoor@dce.ac.in

ACCEPTED MANUSCRIPT

A Novel Fuzzy Energy Based Level Set Method for Medical Image Segmentation

Mahipal Singh Choudhry[1]*, Rajiv Kapoor[2]

Abstract— Segmentation is a very important step in the field of image processing. Noise and intensity inhomogeneity make challenging the segmentation of images, especially for medical images. Fuzzy C-Means (FCM) clustering is one of the most widely used methods in medical image segmentation, but it can't deal effectively with noise and intensity inhomogeneity. Accurate segmentation capability of level set based active contour models make them attractive in medical image analysis but they also fail to perform better when medical images are corrupted by noise. To deal with Gaussian noise and intensity inhomogeneity, a new region-based level set model is proposed by integrating active contour and FCM clustering. In this method, FCM based energy function is used with level set method to overcome local minimum problem of active contour modal. Distance Regularized Level Set Evolution (DRLSE) is used in proposed method to deal with re-initialization problem of traditional level set method. These two modifications in level set modal effectively deal with intensity inhomogeneity of medical image. A Mean filter like spatial term is also utilized with the proposed energy function, which makes this method advantageous for segmenting noisy images. The planned scheme is verified on diverse real medical images and synthetic images, which contain noise as well as intensity inhomogeneity. The proposed method is compared with other state-of-the-art methods in terms of Segmentation Accuracy, Precision, and Recall. Results show that the proposed method offers better performance compared to other latest methods for segmentation of noisy images.

Keywords— **FCM; Level sets; Medical imaging; Segmentation.**

1. Introduction

For medical images, active contour models are very efficient in distinguishing the boundaries. Active contours, referred as snakes, are the models, which are used to extract the objects by making curves evolve by minimizing a defined energy. The energies used in these models are such that they are minimized when the curves evolve at the borders of the required object. Usually, these energies have two main components; internal energy and external energy. Internal energy makes the evolving curve smooth and regularizing it and external energy guides the motion of the curve towards its optimal position. The commonly used external energies are edge based [6], [10], [11] and region based [1], [2], [4], [12]. The active contour models make an automatic search for their minimum energy positions, but sometimes they may settle at a local minimum which makes them less effective. Integration of level sets approaches with edge-based active contour models and then applying them to image processing application was pioneered by Malladi *et al.* [11]. Until then Snake models were implemented based on parameterized approaches which face a lot of complexities. Their proposed method includes an energy term based on edge information into a level set model which normally utilizes curvature motion. A similar approach based on level sets was proposed by Caselles *et al.* [9]. After several years, Caselles *et al.* [6] propose one more approach based on edge information called as Geodesic Active Contours. A gradient flow-based model which is similar to Caselles *et al.* [6] was planned by Satyanand *et al.* [10] in the same year.

Mumford Shah *et al.* [1] proposed a region-based approach to approximate an image using a piecewise smooth model of it. In this approach energy term is minimized when the approximation contains smooth regions as well as a sufficient number of edges that can model the given image. The implementation approach for a particular solution of the Mumford Shah model is given by



ABOUT THE AUTHOR

Sh Mahipal Singh Choudhry is Associate Professor in Department of Electronics & Communication Engineering, Delhi Technological University, India. He received both Bachelor and Master degrees in the field of Electronics & Communication Engineering from Malviya Regional Engineering College, Now Malviya National Institute of Technology, India. His area of interest includes digital image processing, biomedical signal processing.

PUBLIC INTEREST STATEMENT

Image segmentation is the division of an image into its elements/regions and it is a very important aspect of medical image processing used for pathological diagnosis. Segmentation Accuracy is commonly used parameter for performance analysis of any techniques used for image segmentation. Intensity inhomogeneity and noise are the prime hurdles for precise segmentation of medical images. Proposed method deal efficiently with noise and intensity inhomogeneity of medical images.

ACCEPTED MANUSCRIPT

Chang-Vese [4]. In this modal, a piecewise constant approximation of an image is obtained instead of piecewise smooth. The required approximation of the image is achieved is using a level set model developed by *Sethian et al.* [3]. The multiphase extension of [4] is proposed in [12] by Vese *et al.* In the recent years, many advance models are made based on the Chan-Vese [4] model. One of such is proposed by C. Li *et al.* [2] to overcome the intensity inhomogeneity. An energy function based on K-means clustering model is used in the suggested technique. The intensity is modeled using a multiplicative bias field term and a local intensity clustering property is utilized to deal efficiently with inhomogeneity. Chen *et al.* [14] also suggested a method to deal with intensity inhomogeneity almost with the similar approach as in [2] but using an additive bias field. The region-based model depends less on initial level set but edge base models commonly suffer from the problem of level set initialization. B.N. Li *et al.* [7] have used fuzzy c-means clustering's membership function to initialize level set to overcome this problem. They also used the membership function to select the parameters required in the level set model. Cui *et al.* [8] have utilized the local intensity clustering property in forming a new FCM based clustering. L.Tang *et al.* [13] have utilized the concept of introducing fuzziness into the level set approach. They suggested an energy function by integrating FCM_S1&S2 model proposed by Chen *et al.* [16] and level set model proposed by Samson *et al.* [15]. This model successfully segments out images with high noise but most of the methods fail to perform better when medical images have noise and intensity inhomogeneity so in proposed work, a novel fuzzy energy based level set method along with DRLSC and Mean filter based spatial term is proposed for medical image segmentation. The main contributions of the paper are as follows:

- Proposed method effectively deals with intensity inhomogeneity of medical image using FCM based energy function and DRLSC.
- FCM membership function with multiplicative bias field term is used to overcome the problem of local minimum and DRLSC is used to deal with irregularities, developed due to re-initialization during level set evolution.
- Proposed mean filter based spatial term along with FCM based energy function is used to minimize noise effect.
- Proposed method provides superior segmentation results for medical images with intensity inhomogeneity as well as noise and results are compared with other state-of-the-art methods in term of Segmentation Accuracy, Precision, and Recall.

The composition of the research paper is in this way. Section-2 of the research paper deals with background work in this field. The proposed method is discussed in section-3. Section-4 includes results/performance analysis with algorithm and implementation issues. Conclusion, which is followed by references, is explained in section-5.

2. Background

The proposed method is based on Mumford-Shah Model [1] and Two-Phase Chan-Vese Model [4] for image segmentation so these methods are explained in this section with the implementation of Chan-Vese Model for Two-Level Sets.

2.1 Mumford-Shah Model

Mumford and Shah [1] proposed a piece-wise smooth model to segment the image, based on the following assumptions: -

i. The image varies smoothly in the region
ii. The image varies discontinuously across the boundaries between regions.

Let us assume that $\boldsymbol{I}: \boldsymbol{\Omega} \rightarrow \boldsymbol{R}$ is a gray level image, Where Ω represents the image domain. Segmentation of the image is achieved by dividing $\boldsymbol{\Omega}$ into different regions$\{\Omega_1, \Omega_2, \ldots, \Omega_N\}$, those are separated by a contour *C* and represented as

$$\Omega = U_i \Omega_i \cup C \tag{1}$$

To achieve piece-wise smooth approximation of the image ***I,*** the energy function proposed by Mumford-Shah model should be minimized, which is given by

$$F^{MS} = \int_{\Omega}(1-u)^2 dx + \mu \int_{\Omega/C} \nabla u^2 dx + v\,|C| \tag{2}$$

Where, μ and v are fixed parameters. "u" signify the piece-wise soft approximation of the observed image $\boldsymbol{I}$. Value of u is such that it contains smoothly varying values in a region Ω and varies rapidly across the boundaries. The first and second terms in the above model ensure that u follow the above-described properties of the image. The first term makes sure that u is close to the image, *I*. Smoothness of u is controlled by the second term. Term $|C|$ gives the length of the curve to regularize the contour.

ACCEPTED MANUSCRIPT

2.2 Two-Phase Chan-Vese Model

The implementation approach for Mumford-Shah [1] model was proposed by Chan-Vese [4] by taking a special case of the piecewise constant model, where the piecewise smooth model u is replaced by a piecewise constant approximation of it. In this case, energy function given by equation (2), is represented like

$$F^{MC}(\Omega_1, \Omega_2 \ldots\ldots \Omega_N) = \sum_{i=1}^{N} \{\int_{\Omega_i} (I - u_i)^2 \, dx + \int_{\Omega_i} \nabla {u_i}^2 dx + v \left| C_i \right| \} \quad (3)$$

For two-phase Chan-Vese model, two separate regions Ω_1 and Ω_2 in the image are achieved by minimization of above mentioned Chan-Vese model. Regions Ω_1 and $\Omega_2$are separated by a contour C and Ω_1, Ω_2 are inside and outside regions with reference to contour C. It is assumed that the intensities inside these two regions are approximately constant. The fitting energy function used for this purpose is given by,

$$E(C) = E_1(C) + E_2(C) = \int_{\Omega_1} (I - c_1)^2 dx + \int_{\Omega_2} (I - c_2)^2 dx \quad (4)$$

Where, c_1 and c_2 represent the average intensity of an image in the regions Ω_1, Ω_2 respectively.

The fitting function has value as per contour C like

$$E(C) = \begin{cases} E(C_0) = 0; & \text{If } C \text{ is exactly at boundary} \\ E_1(C) > 0, E_2(C) < 0; & \text{If } C \text{ is inside of boundary} \\ E_1(C) < 0, E_2(C) > 0; & \text{If } C \text{ is outside of boundary} \end{cases}$$

By adding regularization terms to the fitting function given by equation (4), energy function gives by Chan-Vese become

$$F(c_1, c_2, C) = \mu \, Length(C) + v \, Area(\Omega) + \lambda_1 \int_{\Omega_1} (I - c_1)^2 dx + \lambda_2 \int_{\Omega_2} (I - c_2)^2 dx \quad (5)$$

Where μ and v are constant parameters whose values can be positive or zero. Values of λ_1, λ_2 affect the weight given to the regions, most of the cases they are taken as unity. *Length (C)* is the length of curve/contour and *Area (Ω)* is the area of the image, *I*. This model will be equivalent to the Mumford-Shah model if $v = 0, \lambda_1 = \lambda_2 = \lambda$ and the constant approximations c_1, c_2 are replaced by piecewise smooth approximations.

2.3 Implementation of Chan-Vese Model for Two-Level Set

If curve/contour $\boldsymbol{C}$ is represented by zero level set $\phi_0(x)$ of level set function $\phi(x)$: $\boldsymbol{\Omega} \longrightarrow \boldsymbol{R}$, two level sets function is defined [3] as

$$\phi(x) = \begin{cases} x \in C; \phi(x) = 0 \\ x \in \Omega_1; \phi(x) > 0 \\ x \in \Omega_2; \phi(x) < 0 \end{cases} \quad (6)$$

And *Length (C)* and *Area (Ω)* are given by

$$Length\,(C) = Length\,\{\phi = 0\} = \int_{\Omega} |\nabla H(\phi(x))| dx = \int_{\Omega} \delta_0(\phi(x)) |\nabla \phi(x)| dx \quad (7)$$

$$Area\,(\Omega) = Area(\phi \geq 0) = \int_{\Omega} H(\phi(x)) dx \quad (8)$$

Where $H(\phi(x)$ represents Heaviside function and $\delta_0(\phi(x))$ represents Dirac delta function. They are defined as

$$H(z) = \begin{cases} 1, if\ z \geq 0 \\ 0, if\ z < 0 \end{cases} \quad (9)$$

$$\delta_0(z) = \frac{dH(z)}{dz} \quad (10)$$

The fitting energy terms for two level sets function become

ACCEPTED MANUSCRIPT

$$E_1(C) = \int_{\Omega_1}(I - c_1)^2 dx = \int_{\phi>0}(I - c_1)^2 dx = \int_{\Omega}(I - c_1)^2 H(\phi(x))dx \quad (11)$$

And $$E_2(C) = \int_{\Omega_2}(I - c_1)^2 dx = \int_{\phi<0}(I - c_2)^2 dx = \int_{\Omega}(I - c_2)^2 (1 - H(\phi(x))dx \quad (12)$$

The Chan-Vese energy function for two level sets become

$$F^{cv}(c_1, c_2, C) = \mu \int_{\Omega} \delta_0(\phi(x))|\nabla\phi(x)|dx + v \int_{\Omega} H\big(\phi(x)\big)dx + \lambda_1 \int_{\Omega}(I - c_1)^2 H(\phi(x))dx + \lambda_2 \int_{\Omega}(I - c_2)^2 (1 - H\big(\phi(x)\big)dx$$

3. Proposed Method

To deal with noise and intensity inhomogeneity of medical image, FCM based energy function is used in the level set method to overcome the local minimum problem of active contour modal. DRLSE is used in the proposed method to deal with the re-initialization problem of traditional level set method. These two modifications in level set modal effectively deal with intensity inhomogeneity of medical image by solving local minimum and re-initialization problems of active contour modal. A Mean filter like spatial term is also utilized with the proposed energy function to minimise noise effect, which makes this method advantageous for segmenting noisy images. Important characteristics of the proposed method are: -

- It is implemented using the level set method.
- A fuzzy-based energy function is derived with consideration of region information and multiplied by a weight factor.
- The intensity of the image is assumed to be modeled by original intensity multiplied by a biased field. The estimation of the biased field is carried out parallel to the evolution of level set using DRLSE, which makes it helpful to deal with inhomogeneity of medical image.
- An improvement to membership of updating function is done by considering the spatial information and second update of membership values by using a spatial term, which results in de-noising of the medical image.

The workflow diagram of proposed method is depicted in Fig.1 as

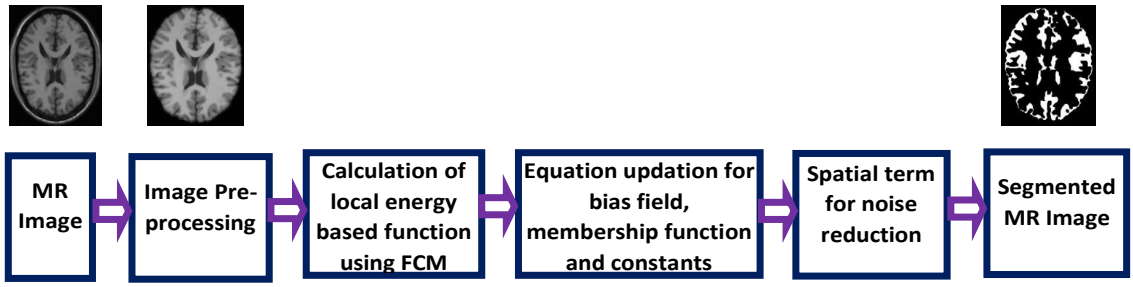


Fig.1: Workflow diagram of proposed method

The proposed method is explained and implemented using following steps: -

i. Image Preprocessing
ii. Calculation of local region based energy function for image by using the Fuzzy C-means concept
iii. Two-Phase Level Set Formulation
iv. Multi-Phase Level Set Formulation
v. Equations Updation
vi. Spatial Term For Reducing Noise Effect

3.1 Image Preprocessing

Edge preserving filtering is used to preserve edge information in images. In proposed research work, the median filter is used for edge-preserving smoothing of medical images. The median filter is a sliding-window spatial filter which replaces the center value with the median of all the pixel values in the window. Usually, the square kernel is used in case of the median filter but

ACCEPTED MANUSCRIPT

other shapes can be also used. A 3x3 window is used in proposed research work for edge-preserving smoothing of medical images.

3.2 Local Region Based Energy Function

The Local region-based model is used to define energy function of proposed model by using the Fuzzy c-means concept. This fuzzy factor assigns each pixel to a particular region based on its membership function. For the medical image, this arrangement of pixel assignment is more suitable instead of hard assignment [16] and this membership function is also utilized to decrease the result of noise. There are two purposes to be achieved using energy function, one is segmentation of image and other is to estimate the bias field. Real world images can be modeled as a multiplicative field, added by noise. This can be represented as

$$I = bJ + n \tag{14}$$

Where I is the observed image, J is the true image, b represents bias filed term, which represents intensity inhomogeneity and n is the additive noise. Salt and Pepper noise, Speckles noise, Blurred noise, Gaussian noise and Poisson noise are commonly affected noises in medical images. Additive Gaussian noise with zero mean and constant variance is targeted in this research work. It is the most common type of noise in medical images results from the contributions of many independent signals [18].

The assumptions made in this model are: -

i. The true image J is assumed as a piecewise constant. i.e. in a given region Ω_i, it takes a constant value c_i.
ii. Bias field b varying very slowly, which implies that in a small neighborhood, b is constant.

In given image, I which is modeled using the equation (14) and based on the above assumptions, the value of bias field, b in a local neighborhood region is considered as constant. i.e. for a circular region, $O_y \subset \{x: |x - y| \leq r\}$ of the image, shown in Fig.2, where r represents the radius of circular region and $y \in \Omega$, the value of bias field is

$$b(x) \approx b(y) \;\; for\, x \in O_y \tag{15}$$

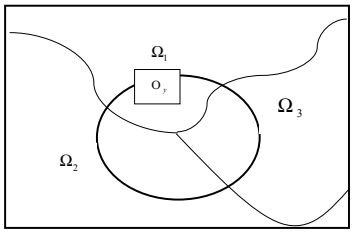


Fig.2: Representation of local region O_y

By using equations (14) and (15), intensity in the local region is modeled as

$$I(x) = b(y)J(x) + n(x) \;\; for\, x \in O_y \tag{16}$$

By using the first postulation about *J(x)* , *I(x)* given by equation (16) is express as

$$I(x) = b(y)C_i + n(x) \;\; for\, x \in O_y \cap \Omega_i \tag{17}$$

The additive Gaussian noise term can be eliminated if pixels, in a region Ω_i, are considered to be part of a Gaussian distribution with mean $m_i = b(y)c_i$

The local region O_y can be divided into N clusters, having centers at $m_i \approx b(y)c_i$, i varying from 1 to N. In this case, energy term of local region by using the FCM is given by

$$F_y = \sum_{i=1}^{N} \int_{O_y \cap \Omega_i} u_i^m(x)\,(I(x) - m_i)^2 dx \tag{18}$$

ACCEPTED MANUSCRIPT

The fuzzy factor m decides the fuzziness and its value is normally taken as 2. In a local region O_y, center of the cluster m_i can be replaced by $b(y)c_i$ and a window function is characterized as

$$W(y-x) \subset \{W(y-x) = 0, x \notin O_y\} \tag{19}$$

By using equations (18) and (19) energy can be written in term of window function as

$$F_y = \sum_{i=1}^{N} \int_{\Omega_i} u_i^m(x) W(y-x)\,(I - b(y)c_i)^2 dx \tag{20}$$

The overall energy by considering the total image domain Ω is given by

$$F(b, c, u, \Omega_1, \Omega_2 \ldots \Omega_N) \triangleq \int \left\{ \sum_{i=1}^{N} \int_{\Omega_i} u_i^m(x) W(y-x)\,(I - b(y)c_i)^2 dx \right\} dy \tag{21}$$

Where, c= $\{c_1, c_2 \ldots . c_N\}$ represents the constants. The selection of Kernel W is flexible. It should take null values outside the local region i.e., $W(z) = 0$ for $|z| > r$ and inside the region, it should have a sum of unity i.e. $\int W(z) = 1$.
$W(z)$ used in this model is

$$W(z) = \begin{cases} \frac{1}{a} e^{-u^2/2\sigma^2}, & for\ |u| \le r \\ 0, & otherwise \end{cases} \tag{22}$$

Selection of the parameter $\boldsymbol{a}$ will depend on the value of $\boldsymbol{\sigma}$ (standard deviation) such that relation $\int W(z) = 1$ is satisfied. The selection of $\boldsymbol{r}$ is one of the crucial factors and the assumptions made on bias field are valid only when the neighborhood is small. The bias field varies faster so the value of $\boldsymbol{r}$ should be small.

3.3 Two-Phase Level Set Formulation

In the two-phase level set model of the proposed model, the image is divided into two regions Ω_1 and Ω_2 using a single level set ϕ. The two regions are defined using the level set as

i. In the region Ω_1, $M_1(\phi) = H(\phi)$
ii. In Ω_2, $M_2(\phi) = \{1 - H(\phi)\}$

Energy function is rewrite using these definitions and equation (21) as

$$F(b, c, u, \phi) \triangleq \int \left\{ \sum_{i=1}^{2} \int u_i^m(x) W(y-x)\,(I - b(y)c_i)^2 M_i\big(\phi(x)\big) dx \right\} dy \tag{23}$$

By exchanging order of integration, equation (23) becomes

$$F(b, c, u, \phi) \triangleq \int \sum_{i=1}^{2} u_i^m(x)\, e_i(x)\, M_i\big(\phi(x)\big) dx \tag{24}$$

Where $e_i(x)$ is defined as

$$e_i(x) = \int W(y-x)\,(I - b(y)c_i)^2 dy \tag{25}$$

$e_i(x)$ can be rearranged in matrix form as

$$e_i(x) = I^2 1_K - 2c_i I(b * K) + c_i^2 (b^2 * K) \tag{26}$$

Where $*$ represents the convolution and

$$1_K = \int W(y-x) dy \tag{27}$$

ACCEPTED MANUSCRIPT

The energy used for level set model is given by

$$F(b,c,u,\phi) = -F(b,c,u,\phi) + vL(\phi) + \mu R_p(\phi) \tag{28}$$

In this equation, $F(b,c,u,\phi)$ is used as data term. $L(\phi)$ and $R(\phi)$ act as regularizing terms which are defined as

$$L(\phi) = \int_\Omega |\nabla H(\phi)| dx \tag{29}$$

And $$R_p(\phi) = \int p\,|\nabla(\phi)|\,dx \tag{30}$$

$L(\phi)$ represents length of contour or zero level set of ϕ, which serves propose of keeping the curve smooth by punishing length of it and term $R_p(\phi)$ is used to avoid reinitialization in level set evolution. Re-initialization is one of the disadvantages of the traditional level set model. During level set evolution, they develop irregularities and leads to numerical errors [6]. To overcome this, formal approach is to stop the evolution by using a signed distance function to redesign it. It is quite tricky to predict suitable time to apply re-initialization. To overcome this conflict, Li et al. [5] planned a term called DRLSC. In level set functions signed distance is maintained by this term. The gradient descent form for evolution of level set function is specified as

$$\frac{\partial \phi}{\partial t} = -\frac{\partial F}{\partial \phi} \tag{31}$$

When energy term $F(b,c,u,\phi)$ is minimized with reference to ϕ (b, c as constants), equation (31) result in

$$\frac{\partial \phi}{\partial t} = \delta(\phi)[(u_1^m e_1 - u_2^m e_2) + v\, div\left\{\frac{\nabla\phi}{|\nabla\phi|}\right\} \mu\, div\{d_p(|\nabla\phi|)\nabla\phi\} \tag{32}$$

Where $$d_p(z) \triangleq \frac{p(s)}{s} \tag{33}$$

To achieve optimal solution using energy function along with level sets function, values of bias field *b* and constants *c* are also updated in a repetitive manner.

3.4 Multi-Phase Level Set Formulation

The two-phase model can be extended to multiphase model, which is utilized for image segmentation using proposed energy function. In this case, k ≥ $\log_2$N level sets are required to solve the N-Phase problem. Ω can be divided into N regions, which are represented by $M_i\{\phi_1(x), \phi_2(x) \dots \dots \dots \phi_k(x)\}$ and are defined as

$$M_i\{\phi_1(x), \phi_2(x) \dots \dots \dots \phi_k(x)\} = \begin{cases} 1, & if\ x \in \Omega_i \\ 0, & Otherwise \end{cases} \tag{34}$$

Energy function for multiphase is given by

$$F(b,c,u,\phi) \triangleq \int \sum_{i=1}^{N} u_i^m(x)\, e_i(x)\, M_i\big(\phi(x)\big) dy \tag{35}$$

Where, $\boldsymbol{\phi} = \{\phi_1, \phi_2 \dots \dots \dots \phi_k\}$ represent the level set vector.

The final energy function with regularizing terms is given by

$$F(b,c,u,\phi) = -F(b,c,u,\phi) + v\sum_{i=1}^{k} L(\phi_i) + \mu \sum_{i=1}^{k} R_p(\phi_i) \tag{36}$$

The gradient descent equations are obtained by minimizing $F(b,c,u,\ \phi)$ with respect to ϕ (b, c as constants) as

$$\frac{\partial \phi_j}{\partial t} = \sum_{i=1}^{N} \frac{\partial M_i(\phi)}{\partial \phi_j}\, u_i^m(x) e_i(x) + \delta(\phi_j)\, v\, div\left\{\frac{\nabla\phi_j}{|\nabla\phi_j|}\right\} \mu\, div\{d_p(|\nabla\phi_j|)\nabla\phi_j\} \tag{37}$$

Where, j = 1, 2……k

ACCEPTED MANUSCRIPT

3.5 Equations Updation

The updating equations for the bias field $\boldsymbol{b}$, member ship function $\boldsymbol{u}$ and constants $\boldsymbol{c}$ are obtained by minimizing $F(b,c,u,\phi)$ with respect to $\boldsymbol{b}$, $\boldsymbol{u}$ or $\boldsymbol{c}$ by keeping other variables constant. To get the update equation for $\boldsymbol{c}$, $F(b,c,u,\phi)$ is minimized with reference to $\boldsymbol{c}$ by keeping $b, u \; and \; \phi$ as constants. For derivative of $F(b,c,u,\phi)$ with respect to c_j by taking $b, u \; and \; \phi$ as constant, all other terms of the summation in equation (35) are independent of c_j, except the j^{th} term so by putting derivative of $F(b,c,u,\phi)$ equal to zero, following relation is obtained

$$\frac{\partial F(b,c,u,\phi)}{\partial c_j} = \int u_j^m(x) \frac{\partial e_j(x)}{\partial c_j} M_j \phi dx = 0 \tag{38}$$

And by taking derivative of $e_i(x)$ (given by equation (25)) with respect to c_j with taking $b, u \; and \; \phi$ as constant

$$\frac{\partial e_j(x)}{\partial c_j} = -2 \int W(y-x)\big(I(x) - b(y)c_j\big)b(y) \tag{39}$$

By using equations (38) and (39)

$$\frac{\partial F(b,c,u,\phi)}{\partial c_j} = \int u_j^m(x) \int W(y-x)\big(I(x) - b(y)c_j\big)\, b(y) M_j(\phi) dx = 0 \tag{40}$$

Or $$\int u_j^m(x) I(x) M_j(\phi) \big(\int W(y-x) b(y) dy\big)\, dx = c_j \int u_j^m(x) M_j(\phi) \big(\int W(y-x) b^2(y) dy\big) dx \tag{41}$$

So $$C_j = \frac{\int (b * W) I M_j(\phi) dx}{\int (b * W) M_j(\phi) dx} \tag{42}$$

Where, $j = 1,2 \ldots . N$

The equation for bias field is updated by minimizing $F(b,c,u,\phi)$ with respect to $\boldsymbol{b}$ by keeping c, u, ϕ as constants. To minimize $F(b,c,u,\phi)$, derivative of $F(b,c,u,\phi)$(given by equation (35)) with respect to $\boldsymbol{b}$ is set equal to zero, which result in the following relation

$$\frac{\partial F(b,c,u,\phi)}{\partial b(y)} = \sum_{i=1}^{k} \int u_i^m(x) \frac{\partial e_i(x)}{\partial b} M_i(\phi) dx = 0 \tag{43}$$

And by taking derivative of $e_i(x)$ (given by equation (25)) with respect to b by taking $c, u \; and \; \phi$ as constant

$$\frac{\partial e_i(x)}{\partial b(y)} = -2W(y-x)(I(x) - b(y)c_i)c_i \tag{44}$$

By using equations (43) and (44)

$$\frac{\partial F(b,c,u,\phi)}{\partial b(y)} = \sum_{i=1}^{k} u_i^m(x) W(y-x)(I(x) - b(y)c_i)c_i \, M_i(\phi) dx = 0 \tag{45}$$

Or $$\int W(y-x) \big\{\sum_{i=1}^{k} u_i^m I(x) M_i(\phi) c_i\big\}\, dx = b(y) \{\int W(y-x) (\sum_{i=1}^{k} u_i^m M_i(\phi) c_i^2) dx\} \tag{46}$$

So $$\hat{b} = \frac{(IJ^{(1)}) * W}{(J^{(2)}) * W} \tag{47}$$

Where $$J^{(1)} = \sum_{i=1}^{N} M_i(\phi) c_i u_i^m \tag{48}$$

And $$J^{(2)} = \sum_{i=1}^{N} M_i(\phi) c_i^2 u_i^m \tag{49}$$

To get updating equation for u_j, energy term is solved using Lagrange multiplier as

ACCEPTED MANUSCRIPT

$$F_m = F + \lambda\{1 - \sum_{i=1}^{k} u_i^m(x)\} \quad (50)$$

And $\sum_{i=1}^{m} u_i^m(x) = 1\,, \forall\, x$ (51)

These equations are updated by considering derivative of $\boldsymbol{F_m}$ with reference to u_j (with c, b, ϕ as constants) and equating it to zero. The estimated value of u_j can be found as

$$\frac{\partial F_m}{\partial u_j(x)} = m u_j^{m-1}(x) e_j(x) - \lambda = 0 \quad (52)$$

$$\Rightarrow \hat{u}_j = \left[\lambda / \{m u_j^{m-1}(x) e_j(x)\}\right]^{\frac{1}{m-1}} \quad (53)$$

And $\sum_{i=1}^{k} u_i^m(x) = 1, \forall x$ (54)

By using equations (53) and (54)

$$\sum_{i=1}^{k} [\lambda / \{m u_i^{m-1}(x)\, e_i(x)\}]^{m/(m-1)} = 1 \quad (55)$$

So $\lambda = [\sum_{i=1}^{k} \{m u_i^{m-1}(x) e_i(x)\}^{-1/(m-1)}]^{-(m-1)}$ (56)

By substituting value of λ in equation (53), gives

$$\hat{u}_j = \frac{\{u_j^{m-1}(x) e_j(x)\}^{-1/(m-1)}}{\sum_{i=1}^{k} \{u_i^{m-1}(x) e_i(x)\}^{-1/(m-1)}} \quad (57)$$

3.6 Spatial Term For Reducing Noise Effect

A spatial term for fuzzy clustering using membership function is used in the level set formulation of the proposed method to subjugate noise effect. At every step along with bias filed, constants, membership functions, the spatial term is also calculated. In proposed method, spatial term is average of membership values of neighboring pixels and given as

$$h_i(x) = \sum_{x_r \in N_R(x)} u_i(x_r) \quad (58)$$

By taking this extra spatial term into consideration, a pixel membership value is decided by the neighboring pixels. i.e., even if the center pixel is noisy its effect can be truncated. Using this spatial term, the membership function is updated as

$$u_i(x) = \frac{u_i^p(x) h_i^q(x)}{\sum_{j=1}^{k} u_j^p(x) h_j^q(x)} \quad (59)$$

Where *p* and *q* represent the weight given to each term. If noisy is heavy, more weight is given to the spatial term by considering large values for *q*. Median filter, Wiener filter, and Gaussian filter are also tested in the proposed method in place of Mean filter based spatial term. The median filter performs better for removing Salt and Pepper noise and Poisson noise, Wiener filter performs for removing Speckles noise and Gaussian filter for the Blurred noise. But in case of additive Gaussian noise with zero mean and constant variance, mean filter based spatial term gives best results.

4. Results

The proposed method is verified with numerous synthetic images and real medical images and result are compared with some of the states of the art techniques. A CT scan image of the heart is taken from Cornell university database (www.via.cornell.edu/database) and MR images are taken from database (www.humanconnectome.org) of Connectome Coordination Facility (CCF). Fig.3 display corresponding results against evaluation of proposed method on a CT scan image of heart as

ACCEPTED MANUSCRIPT

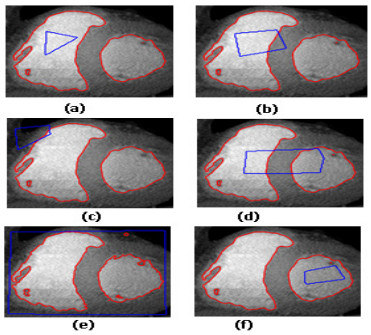


Fig.3: Contour Evaluation using Different Types of Initialization

Initial contour in Fig. 3(a) and Fig. 3(f) is inside of the region of interest and in Fig. 3(e), it is completely outside of the region. In all case, the final contour is same and results are independent of initialization of initial contour. Here proposed method is used without the spatial term. If the spatial term is also included in the proposed method, results are same since these images are without noise.

The original CT scan image, final segmented image, bias field and corrected image using the bias filed are given in Fig.4 as

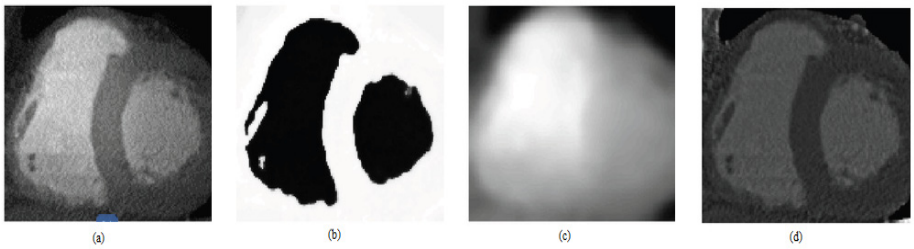


Fig.4: (a) Original CT scan image (b) Final segmented image (c) bias field (d) Corrected image using the bias filed

In Fig.5-7, Brain MRI images obtained from web dataset of CCF are used. The MRI data used is T1 normal MRI of 1mm thickness with noise density of 9% and intensity non-uniformity of 40%.

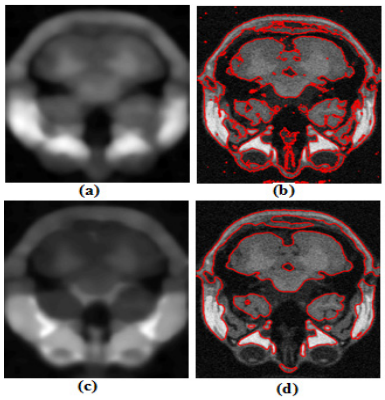


Fig.5 Results of the proposed method and C. Li *et al* [2] on a noisy MRI image. (a, b) Bias field & Final contour using Li's method (c, d) Bias field & Final contour using the proposed method.

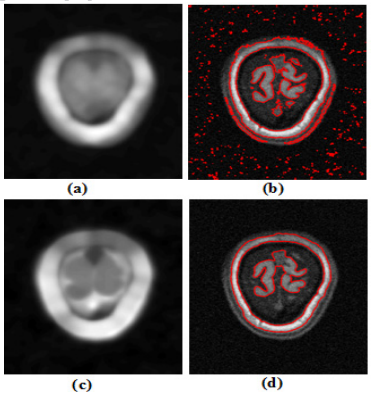


Fig.6 Results of the proposed method and B.N. Li *et al.* [7] on a noisy MRI image. (a, b) Bias field & Final contour using B.N Li's method (c, d) Bias field & Final contour using the proposed method.

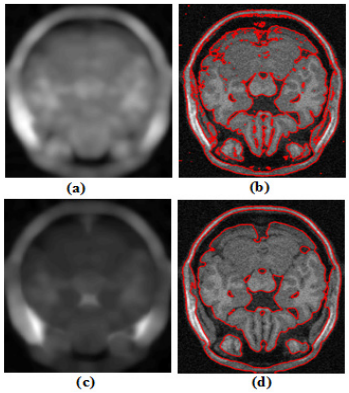


Fig.7 Results of the proposed method and W. Cui *et al.* [8] on a noisy MRI image. (a, b) Bias field & Final contour using W. Cui method (c, d) Bias field & Final contour using the proposed method.

It can be observed from the images that final contour obtained by other methods is affected by the noise and in case of the proposed method, the noise effect is either almost zero or very less. The proposed method gives superior results both in terms of bias field and final contour.

ACCEPTED MANUSCRIPT

Results of the proposed method are compared with results of other methods for many synthesis images corrupted by different Gaussian noise (noise density,$\sigma = 0\ to\ 0.1$) and corresponding results of a synthesis image, added with Gaussian noise density (σ) of 0.02, are publicized in Fig.8.

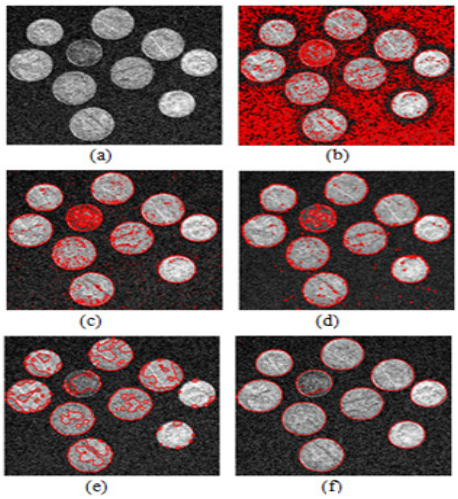


Fig.8 (a) Noisy Image with Gaussian Noise Density of σ = 0.02. Results using (b) C. Li *et al* [2] (c) W. Cui *et al.* [8] (d) Y. Chen *et al.* [14] (e) B.N. Li *et al*. [7] (f) Proposed Method

The proposed method is compared with other latest methods on the basis of Segmentation Accuracy (SA) for same synthetic image corrupted by different Gaussian noise (noise density,$\sigma = 0\ to\ 0.1$). The values obtained by experimental evaluation are shown in table.1.SA is defined as [17]

$$SA = \frac{Number\ of\ Correctly\ Classified\ Pixels}{Total\ Number\ of\ Pixels} \tag{60}$$

Method	***Noise Density***					
	0	0.02	0.04	0.06	0.08	0.10
C. Li *et al* [2]	92.81	88.26	86.18	83.65	78.95	76.56
W. Cui *et al.* [8]	95.64	92.44	89.92	88.06	85.22	82.47
Y. Chen *et al.* [14]	92.78	89.57	85.32	82.48	78.12	77.21
B.N. Li *et al*.[7]	96.42	93.30	88.74	86.15	84.74	85.13
Proposed Method	99.85	98.23	96.28	94.76	93.42	92.38

Table 1: **Segmentation Accuracy (SA) for Synthetic Image Corrupted by Gaussian Noise**

ACCEPTED MANUSCRIPT

C. Li *et al* [2]: A region-based method with K-means clustering based energy function for image segmentation.
B.N. Li *et al* [7]: A level set method with fuzzy c-means clustering's membership function for image segmentation.
W. Cui *et al.* [8]: Localized FCM clustering with spatial information for medical image segmentation and bias field estimation.
Y. Chen *et al.* [14]: An improved level set method using additive bias field for brain MR images segmentation.

It can be observed that in case of segmentation of synthesis image, the proposed method has the minimum effect of noise compared to other methods. Precision and Recall curves are also commonly used for quantitative evaluation of image segmentation methods. In case of classification, Precision and Recall are referred as Positive Predictive Value (PPV) and True Positive Rate (TPR) or sensitivity respectively. Precision and Recall are defined as

$$Precision = \frac{TP}{(TP+FP)} \tag{61}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{62}$$

Where TP, FP, and FN are True Positive, False Positive and False Positive respectively as per Fig.9 (Confusion Matrix)

		True Condition	
		P	**N**
Predicted Condition	**P**	**TP**	**FP**
	N	**FN**	**TN**

Fig.9 Confusion Matrix

Precision and Recall curves of the proposed method along with state-of-the-art methods are plotted in Fig.10 for synthesis image corrupted with Gaussian noise (σ) of 0.02.

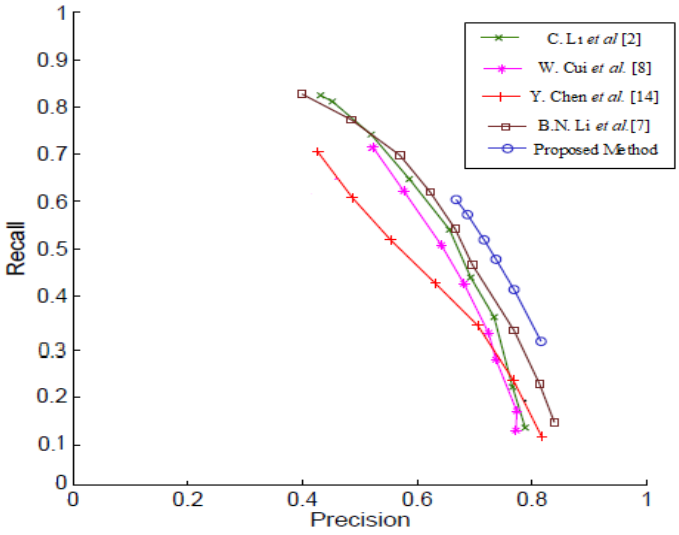


Fig.10 Precision and Recall curves for synthesis image with Gaussian noise density of 0.02.

ACCEPTED MANUSCRIPT

The precision and recall curves also verify that proposed method produces better segmentation. Larger recall and precision values in case of other methods come at the cost of undesirable over segmentation or under segmentation.

The proposed method is verified with different types of medical images like CT scan images of heart and brain MR images. The results show that proposed method efficiently deal with noise as well as intensity inhomogeneity problem of synthesis and real time medical images.

5. Conclusion

Results show that in the proposed method, contour evaluation results are independent of the initialization of initial contour and final contour obtained by the proposed method is less affected by noise in comparison to other methods. Proposed method results in smooth contour even in existence of noise and non-uniformity. When the proposed method is compared with other methods in term of Segmentation Accuracy of synthesis image corrupted by Gaussian Noise along with precision and recall curves, it shows better results. Median filter, Wiener filter, and Gaussian filter are also tested in the proposed method in place of Mean filter based spatial term. In case of additive Gaussian noise with zero mean and constant variance, mean filter based spatial term gives best results. The proposed method has the better segmentation of medical images corrupted by both intensity inhomogeneity as well as noise so proposed method can be utilized for medical applications, where high quality and precise segmentation are required. The proposed method is computationally complex since multiple similar convolutions are used repeatedly. As a future work, complexity can be reduced by using different types of level set methods for implementation of active contours and energy function using different FCM variations.

Acknowledgment

Thank Dr. Dinesh Kumar Vishwakarma for his valuable comments and support. http://www.dtu.ac.in/Web/Departments/Electronics/faculty/dkvishwakarma.php

Funding

Not received any funding for this research.

Author details

Mahipal Singh Choudhry[1]
E-mail: msc_1976@yahoo.com
Rajiv Kapoor[2]
E-mail: rajivkapoor@dce.ac.in
[1]Department of Electronics & Communication Engineering, Delhi Technological University, Delhi, India.
[2]Ambedkar Institute of Advanced Communication Technologies & Research, Delhi, India.

Citation information

Cite this article as A Novel Fuzzy Energy Based Level Set Method for Medical Image Segmentation, Mahipal Singh Choudhry, Rajiv Kapoor, *Cogent Engineering* (2016)

References

Mumford and J. Shah (1989), Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems**,** *Communications on Pure and Applied Mathematics,* vol. XLII, 577-685.
http://dx.doi.org/10.1002/cpa.3160420503

C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas and John C. Gore (2011), A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities with Application to MRI, *IEEE Transactions on Image Processing,* Vol. 20, No. 7.
http://dx.doi.org/10.1109/TIP.2011.2146190

S. Osher and J. Sethian (1988), Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations, *Journal of Computational Physics*, Vol. 79, No. 1, pp. 12–49.
http://dx.doi.org/ 10.1016/0021-9991(88)90002-2.

T. Chan and L. Vese (2001), Active contours without edges, *IEEE Transactions on Image Processing*, Vol. 10, No. 2, pp. 266–277.
http://dx.doI.org/ 10.1109/83.902291

C. Li, C. Xu, C. Gui, and M. D. Fox (2010), Distance Regularized Level Set Evolution and Its Application to Image Segmentation, *IEEE Transactions on Image Processing,* Vol. 19, No. 12.
http://dx.doi.org/10.1109/TIP.2010.2069690

V. Caselles and R. Kimmel (1997), Geodesic Active Contours, *International Journal of Computer Vision,* 22(1), 61–79.
http://dx.doi.org/ 10.1023/A: 1007979827043

B. N. Li, C. Chui, S. Chang and S. H. Ong (2011), Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation, *Computers in Biology and Medicine,* 41(1):1–10.
http://dx.doi.org/10.1016/j.compbiomed.2010.10.007

W. Cui, Y. Wang, Y. Fan, Y. Feng, and T. Lei (2013), Localized FCM Clustering with Spatial Information for Medical Image Segmentation and Bias Field Estimation, *International Journal of Biomedical Imaging, V*olume 2013, Article ID 930301.
http://dx.doi.org/10.1155/2013/930301

V. Caselles, F. Catte, T. Coll, and F. Dibos (1993), A geometric model for active contours in image processing, *Numer. Math*., Vol. 66, No. 1, pp. 1–31.
http://dx.doi.org/10.1007/BF01385685

S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi (1995), Gradient flows and geometric active contour models, *Proceeding, 5th Int. Conf. Computer Vision*, pp. 810–815.
http://dx.doi.org/10.1109/ICCV.1995.466855

R. Malladi, J. A. Sethian, and B. C. Vemuri (1995), Shape modeling with front propagation: A level set approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 2, pp. 158–175.
http://dx.doi.org/10.1109/34.368173

L. A. Vese and T. F. Chan (2002), A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model, *International Journal of Computer Vision, 4*0(3), 271–293.

ACCEPTED MANUSCRIPT

http://dx.doi.org/10.1023/A: 1020874308076
L. Tang (2014), A Variational Level Set Model Combined with FCMS for Image Clustering Segmentation, *Mathematical Problems in Engineering, V*ol. 2014, Article ID 145780.
http://dx.doi.org/10.1155/2014/145780
Y. Chen, J. Zhang and J. Macione (2009), An improved level set method for brain MR images segmentation and bias correction, *Computerized Medical Imaging and Graphics, V*ol 33, Issue 7, Pages 510–519.
http://dx.doi.org/10.1016/j.compmedimag.2009.04.009
C. Samson, L. Blanc-F´eraud, G. Aubert and J. Zerubia (2000), Level set modal for image classification, *International Journal of Computer Vision*, Vol. 40, No. 3, pp. 187–197.
http://dx.doi.org/10.1023/A:1008183109594
S. Chen and D. Zhang (2004), Robust Image Segmentation Using FCM with Spatial Constraints Based on New Kernel-Induced Distance Measure, *IEEE Transaction. on System., man and Cybernetics*—Part b: Cybernetics, Vol. 34, No. 4.
http://dx.doi.org/10.1109/TSMCB.2004.831165
Mohamed N. Ahmed, Sameh M. Yamany, Nevin Mohamed, Aly A. Farag, and Thomas Moriarty (2002), A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data, *IEEE Transactions on Medical Imaging*, Vol. 21, No. 3.
http://dx.doi.org/10.1109/42.996338.
Pierre Gravel, Gilles Beaudoin and Jacques A. De Guise (2004), A Method for Modeling Noise in Medical Images, *IEEE Transactions on Medical Imaging*, Vol. 23, No. 10.
http://dx.doi.org/10.1109/10.1109.
Mahipal Singh Choudhry and Rajiv Kapoor (2016), Performance Analysis of Fuzzy C-Means Clustering Methods for MR Image Segmentation, Procedia Computer Science, Vol 89,Pages 749-758.
http://dx.doi.org/10.1016/j.proces.2016.06.052.
Pranali A.Tembhekar, M. N. Thakare and S. A. Dhande (2015),Spatial Fuzzy Clustering With Level Set Method for MRI Image, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 7.
http://dx.doi.org/10.17148/IJARCCE.2015.4756
Jian Tang and Xiaoliang Jiang (2017), A Variational Level Set Approach Based on Local Entropy for Image Segmentation and Bias Field Correction, Computational and Mathematical Methods in Medicine, Volume 2017, Article ID 9174275.
http://dx.doi.org/10.1155/2017/9174275
Ankita S. Pounikar and Sheeja S. Suresh (2015), Spatial K Fuzzy and level set approach for image segmentation in brain MRI for tumor detection, IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 5, Issue 2.
http://dx.doi.org/ 10.9790/4200-05224548.
Maryam Rastgarpour and Jamshid Shanbezadeh (2014), A New Kernel-Based Fuzzy Level Set Method for Automated Segmentation of Medical Images in the Presence of Intensity Inhomogeneity, Computational and Mathematical Methods in Medicine, Computational and Mathematical Methods in Medicine.
http://dx.doi.org/10.1155/2014/978373

ACCEPTED MANUSCRIPT

Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

A STUDY ON INCREMENTAL MINING OF FREQUENT PATTERNS

M. Dutta Borah[1,*], R. Jindal[2]

[1]Inderprastha Engineering College,U.P., India

[2]Delhi Technological University, New Delhi, India

Received: 01 March 2018 / Accepted: 29 April 2018 / Published online: 01 May 2018

ABSTRACT

Data generated from both the offline and online sources are incremental in nature. Changes in the underlying database occur due to the incremental data. Mining frequent patterns are costly in changing databases, since it requires scanning the database from the start. Thus, mining of growing databases has been a great concern. To mine the growing databases, a new Data Mining technique called Incremental Mining has emerged. The Incremental Mining uses previous mining result to get the desired knowledge by reducing mining costs in terms of time and space. This state of the art paper focuses on Incremental Mining approaches and identifies suitable approaches which are the need of real world problem.

Keywords: Data Mining, Frequent Pattern, Incremental Mining, Frequent Pattern Minung, High Utility Mining, Constraint Mining.

Author Correspondence, e-mail: malaya.dutta@ipec.org.in

doi: http://dx.doi.org/10.4314/jfas.v10i2.8

1. INTRODUCTION

For every organization to be in the right place at right times are the crucial factors to ensure right decision from large amount of data [1]. To deal with the large amount of data stored in the database, Data Mining (DM) has been introduced. A widely accepted definition of DM is: "*It is the process of discovering interesting knowledge from large amounts of data stored in*

Journal of Fundamental and Applied Sciences is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. Libraries Resource Directory. We are listed under Research Associations category.

databases, data warehouses, or other information repositories [2]". Various DM techniques such as Frequent Pattern Mining, Incremental Mining of Frequent Patterns**,** Association Rule Mining, Classification, Prediction, Clustering, and Neural Networks etc. are discussed in various literatures.

Frequent Pattern Mining (FPM) is an important research topic in the field of DM to discover interesting patterns in databases. A pattern is said to be *interesting pattern* if it appears at least as frequently as a predetermined minimum support threshold. Mining frequent patterns is costly during changes of the underlying database. Most of the traditional approaches need repetitive scanning of the database as well as generating huge number of candidate sets [3].

The **Incremental Mining (IM)** is an important research topic in the field of DM, uses previous mining result to get the desired information by reducing mining costs in terms of time and space [4]. Incremental Mining is broadly defined as the way of extracting knowledge from dynamic dataset, where new instances of transactions keep adding and the obsolete instances keep deleting from the mining process (Refer Figure 1 (a) & (b)).

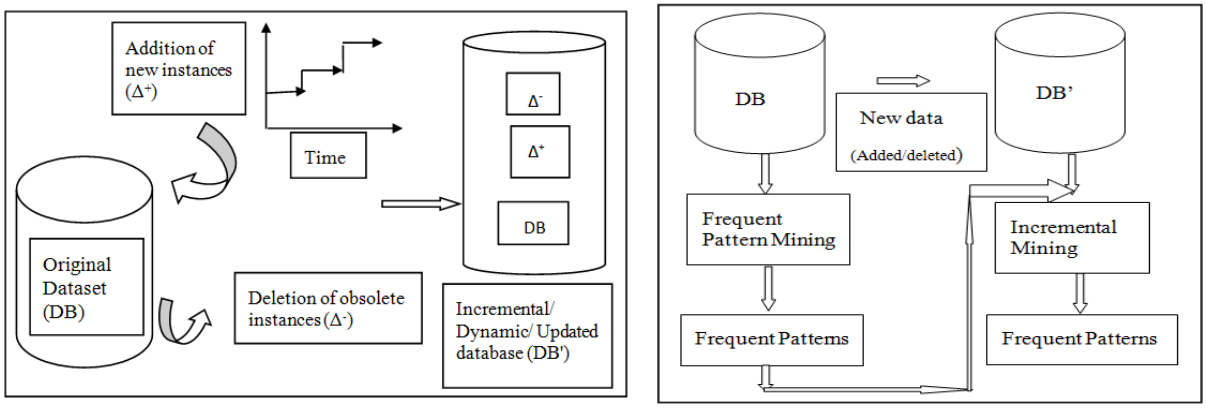


Fig.1. (a) Incremenmtal Database (b) Incremental Mining of Frequent Patterns

The important issue is that how to utilize the mining result to efficiently reflect these changes. The motivation of this work is to find research trends of Incremental Mining of frequent patterns with respect to Research Questions (RQs) listed below.

- **RQ#1:** What is the need for Incremental Mining?
- **RQ#2:** What are the published literature/approaches for Incremental mining of frequent patterns?

- **RQ#3:** Can they be classified according to functionality?
- **RQ#4:** Are they able to measure the semantic significance of an item as per user requirements?
- **RQ#5:** Does they satisfy users' constraints?

The rest of the paper is organized into 6 sections: The next section, Section 2 Review of Incremental Mining approaches are discussed, in Section 3 Key Resarch Findings are presented, Section 4 discusses the Goal-realted issues, Section-5 presented Application of Incremental Approaches, Section 6 we summarize the overall reviewed approaches as Conclusion and in Section 7 References we have included in this work are presented.

2. REVIEW OF INCREMENTAL MINING APPROACHES

Incremental data can be defined as dynamic data that changes with time. In real world applications, underlying database consists of Incremental data that keeps changing and requires continual updating. Mining frequent patterns from such type of changing/growing databases is costly, as most of the approaches need repetitive scanning of the databases as well as generating huge number of candidate sets [3].

Frequent pattern mining is an important data mining technique that can be utilized to discover the frequent patterns from changing/growing database [5]. It is important to take into account the functionality (what kind of patterns) and performance (how to mine frequent patterns). On the basis of this concept, the research works are divided into the following three categories chronologically:

- Apriori based Incremental Mining
- Partition based Incremental Mining and
- Pattern Growth based Incremental Mining
- High Utility based Incremental Mining
- Constraint based Incremental Mining

2.1. Apriori based Incremental Mining

This section describes a candidate set generation process based on Apriori based Incremental Mining approach. We have considered those approaches that are widely used by the research

community as listed below.

Fast UPdate (FUP) (1996) [6]**,** is the first Apriori based Incremental Mining algorithm is proposed for the efficient maintenance of discovering association rules in large databases when new transaction data are added to a transaction database. It generates less number of candidate keys as compared to Apriori, but deletion operation is not incorporated in this algorithm.

Fast UPdate2 (FUP2) (1997) [7] is an extended version of FUP, which handles the association rules in both Insertion and Deletion operation. This algorithm assumes that the large itemsets and their Support Count in the database came from the results of previously stored mining activities. It is more desirable to store counts rather than the association rules as the association rules can be calculated from these counts efficiently. The focus of this algorithm is that, the transaction to be deleted is a small part of the database activity.

Update With Early Pruning (UWEP) (1999) [8]**,** is another Apriori based algorithm that updates association rules with early pruning. It uses a dynamic look ahead strategy in updating the existing large itemsets. This algorithm detects and removes those items that will no longer remain large after the contribution of the new set of transactions. The limitation is that it does not allow changes in minimum support value.

PELICAN (2001) [9] algorithm is similar to FUP2 algorithm which is based on a lattice decomposition strategy to find out the maximum frequent/ large itemset when the database is updated. Apart from that it allows changes in minimum support value and the generation of candidate sets is less, compared to FUP algorithm. Limitation of the algorithm is that, it considers only maximum frequent itemset in mining process.

2.2. Partition based Incremental Mining

The limitation of the Apriori based Incremental Mining Algorithms are that they require multiple numbers of scans as well as generate large candidate sets. To cope up with this problem **[10]** developed a partition-based algorithm known as **Sliding Window Filtering (SWF)** algorithm. In this algorithm, a transaction database, 'D' is divided into 'n' number of partitions i.e. {P1, P2---Pn} and it employs filtering threshold in each partition to deal with the candidate itemset generation. The Cumulative Filter (CF) produced in processing each partition constitutes the key component to realize the incremental mining. If 'L' is a frequent itemset in

'D', then 'L' must be frequent at least in one of the 'n' number of partitions.

An extended version of SWF is proposed by [11]**.** Two algorithms have been proposed by reusing previous mining task for generating frequent patterns, namely SWF with Frequent Itemset (FI SWF) and SWF with Candidate Itemset (CI SWF).

This approach is faster and generates fewer candidates as compared to SWF and Apriori based algorithms. The limitation of this algorithm is that it does not take into consideration the situation that the dataset to be Inserted or Deleted is either too large or too small.

2.3. Tree/Pattern Growth based Incremental Mining

The limitations of the Apriori based Incremental Mining Algorithms and Partition based Incremental Mining Algorithms are:

- Multiple no. of scans required
- Do not allow changes in minimum support value
- Large candidate sets
- Assumption that data resides only on primary memory is not possible in most of the real-world applications

Taking into account of the above mentioned four limitations, Tree/Pattern Growth based Incremental Mining Algorithms has been introduced by the researchers.

Compressed and Arranged Transaction Sequences Tree (CATS Tree) (2003) [12] is an extension of FP tree which improves storage compression and allow frequent pattern mining without generation of candidate itemsets. The procedure of tree construction of CATS tree is as follows:

- Initialize CATS tree as Empty tree
- Calculate its itemset frequency. Frequency of Parent node always greater than sum of children node.
- The all itemset appears in the first transaction added to the initial tree. During the addition of next transactions, common items are extracted and merge with the existing tree if Parent node frequency > sum of Children node frequency. Similar procedure applicable for rest descendent. Remaining transactions are inserted as a new branch at the parent node.

Another variation of CATS is **FELINE** (Frequent/Large pattern mining with CATS tree) that generates frequent patterns without generating candidate set. It follows divide and conquer strategy. It partitions the dataset according to the transactions it has.

Canonical order Tree (CanTree) (2007) [13] is a tree where the nodes are arranged according to Canonical order. It is suitable as compared to CATS is that no adjustment of nodes is required in case of modification of the tree.

Fast updated Frequent Pattern Tree (FUFP) (2006) [14] is similar to FP-Tree except that the links between parent nodes and their child nodes are bi directional which helps fasten the process of deletion in the maintenance process.

Incremenatl Mining Binary Tree (IMBT) **(2009)** [15] approach does not require predetermining the minimum support threshold. This algorithm efficiently enumerates the support of each item set.

Pre-Frequent Pattern(Pre-FP) algorithm **(2010)** [16] based on the theory of pre-large itemsets, predefined support threshold named as "upper" and "lower". When the new transaction is added the frequent item do not transform into infrequent ones or vise versa, but the items are put into prelarge itemset first.

Item-Itemset (I-Is) (2012) [17] is based on IMBT. The item is placed in the tree depending upon its frequency. Original dataset is scanned only scans and frequency of each item and itemset is calculated and I-Is tree is constructed. When incremental data arrives, the I-Is tree is updated.

Tree for Incremental Mining of Frequent Pattern (TIMFP) (2016) [18] is a binary search-based tree structure where building the tree is unaffected by changes in item frequency as it uses binary search tree property to build the tree. This work is suitable for both interactive and incremental mining.

2.4. High Utility Frequent Pattern Incremental Mining

The approaches discussed in the above section was well defined on binary domain with equal measure of interestingness of items. In real world scienerio it is not always possible to have equal measures. Interestingness measures of items can be calculated in context of Unit cost, Profit, Priority, Distance, Credit e.g. Cost of items in a transactional database, importance of the

traversing different web pages as per the priority basis etc.

Several algorithms have been proposed by researchers to reflect the Interestingness measures of items. Few are listed below:

Weighted Frequent Itemset Mining (WFIM) (2005) [19]is the first FP tree-based algorithm that reflects the relative importance (weight) of an item. Items are assigned with fixed weight in ascending order in the tree.

Weighted Interesting Patterns (WIP) (2006) [20] is a remarkable algorithm which introduces a new measure, weight-confidence to mine correlated patterns with weight affinity. Researcher have pointed out that weighted frequency of an item does not have the downward closure property.

Weighted Itemsets (WIT) (2013) [21] is used to fast mining of Frequent Weighted Itemsets (FWI) from transaction databases. In such framework, two tree structures i.e. Incremental Weighted Frequent Pattern Tree based on Weight Ascending order (IWFPTWA) & Incremental Weighted Frequent Pattern Tree based on Frequency Descending order (IWFPTFD) and two algorithms IWFPWA & IWFPFD are proposed by [22].

Some of the remarkable research in high utility incremental mining includes Utility Pattern Growth+ (UP-Growth+) and Utility Pattern Tree (UP tree) [23], Fast High-Utility Miner (FHM) [24], Maximum Utility Growth (MU-Growth, MU-Strategy1, MU-Strategy2) and Maximum Item Quantity Tree (MIQ-Tree) [25].

High Utility Mining (HUM) (2017) [26]**:** This approach is utility is profit and quantiuty and applied in education domain dataset. This is suitable to calculate the incremental semester wise performance of the students.

Mine High Average-Utility Patterns with Multiple Minimum Average-Utility Thresholds (HAUIM-MMAU) (2018) [27]**:** This approach uses average-utility list structure with multiple utility threshold that reduces mining time instead of coverarging the maximum ulility pattern. This approach works on the build once and mine property.

2.5. Constraint based Incremental Mining

Constraint based Incmemental Mining deals with the mining of patterns what exactly a user wants to extract instead of generating all the patterns. Following are some constraint based

incremental mining algorithm.

Gap with Index Apriori (GwI-Apriori) (2009) [28]: uses gap constraint and level-wise approach to mine the frequent patterns. The limitation of this approach is that it spends much time in disk I/O hence not suitable for mine large amount of data.

RegularMine (2010) [29] uses an extension of itemsets called regular to investigate subset, superset or proper subset of items present or not. This approach produces user required frequent patterns from a large amount of closed itemsets.

Maximum Constraint Rare Pattern Tree (MCRP-Tree) (2015) [30] is FP tree-based approach to mine rare itemset and genereate association among them. Rare items are the item that occurs uncommonly in the dataset.

Constraint based Incremental Frequent Pattern Mining (CIFMine) (2015) [31]is an incremental constraint mining approach where dataset filtering techniques are used to reduce the size of the dataset. The main constraints discussed in this work are pattern, pattern length user defined minimum support threshold. This work is suitable for buid once and mine many constraint patterns from dynamic dataset.

NegI-NSP (Negative Sequential Patterns) (2017) [32]is a recent work which introduces strict constraints to solve a series of consequent problems. This approach starts with calculating negative sequnces then algorithm **NegI-NSP** applied to mine the useful patterns.

NOSEP (Nonoverlapping Sequence Pattern Mining With Gap Constraints) (2018) [33] is a most recent work focuses on maximally utilised pattern discovery using gap constraint. This research uses Nettree data structure to build and mine the pattern tree.

3. KEY RESEARCH FINDINGS

3.1 Apriori based Incremental Mining Algorithm

- These algorithms are mostly suitable for transactional databases
- Iteration based algorithm
- Huge number of Candidate set generation both in original as well as incremental portion which is undesirable in Incremental Mining process
- There is no concept of pruning particularly in FastUPdate based algorithms so generates a large number of candidate sets. But UWEP and PELICAN work with the concept of

early pruning

- In data base updatation, rescans are required both in the original and incremental databases
- Minimum Support value is fixed in both FastUPdate based and UWEP algorithms where as PELICAN allows changes in minimum support
- Overall assumption is that computation of all data takes place main memory only
- Does not consider the situation that transaction dataset to be inserted or be deleted too large or too small.

3.2. Partition based Incremental Mining Algorithms

- These algorithms are suitable for transactions as well as temporal databases. Particularly time –variant transaction databases.
- Partition based algorithm and employs filtering threshold in each partition to deal with the candidate itemset generation.
- Reduces the number of scans by using previous resultant candidate sets. Only one scan is required in incremenatal portion.
- Faster mining and less space required as compared to fast update-based algorithms because no need to keep all large itemsets in the main memory. Only one has to keep the resultant candidate set for further incremental mining process.
- Does not consider the situation that transaction dataset to be inserted or deleted is too large or too small.

3.3. Tree/Pattern Growth based Incremental Mining

- The tree is built without generating candidate sets
- Cope up of different support value without rebuilding the tree
- Apriori based incremental mining algorithm need multiple passes for an incremental mining process where as most of pattern growth algorithm need only single pass.
- Concurrent deletion is possible
- Faster mining and less space required as because uses the concept “build once and mine many”.
- The common assumption is, each item in a database is equal in weight

3.4. High Utility Incremental Frequent Pattern Mining

- The tree captures items along with weight in transaction database and arranges tree nodes according to normalized weight in ascending order.
- It is easily manageable when performed Insertion, Deletion and Update and mining operations.
- It fulfills "Build once and Mine many" which is suitable for incremental as well as interactive mining.

3.5. Constraint based Incremental Mining

- Focuses on the constraints and to specify the desired properties of the patterns to be mined that are likely to be of interest to the end user
- Identifies different constraints
- Suitable for semi structured/unstructured data

4. GOAL-REALATED ISSUES

This state-of art work presents a systematic way to address the frequent pattern mining problem using incremental frequent pattern mining. The following table shows goal related issues with respect to the RQs listed in the introduction section.

Table 1. Goal related issues of Imcremental Mining w.r.to RQS

	RQs	Goal-related issues addresses w.r.to RQs
1	What is the need for Incremental Mining?	Cost of Frequent Pattern Mining
2	What are the published literature/ approaches for Incremental mining of frequent patterns?	Review of Incremental Mining Approaches
3	Can they be classified according to functionality?	Problems in terms of Computational and Algorithmic Development
4	Are they able to measure the semantic significance of an item as per user requirements?	Importance of Semantic Significance
5	Do they satisfy users' constraints?	Importance of Constraints

The following goal-related issues have been addressed:

- **Cost of Frequent Pattern Mining**: Mining frequent patterns are costly during changes of the underlying database. Most of the traditional approaches need repetitive scanning of the database as well as generating huge number of candidate sets. In such type of situation, the important issue is how to utilize the mining result efficiently to reflect these changes. Thus, mining of growing databases is an interesting area of research in Data Mining.
- **Problems in terms of Computational and Algorithmic Development**: This causes Creation of duplicate node, Greater computation time required for Searching common itemsets, Merging and finding a suitable path, Consumption of large amounts of memory to construct the tree and greater mining time etc. It is important to develop efficient approaches to enhance the performance of mining.
- **Importance of Semantic Significance**: In Frequent pattern Mining, the common assumption is that each item in a database is equal in weight. As a matter of fact, in real life applications, it is not always possible to treat all patterns in an equal or binary way. Utility/weight associated with items is important in real life applications and it provides Semantic Significance of items presents in a database.
- **Importance of Constraints**: It is not always suitable for end users that all generated patterns are equally important. Constrains are important in cases where end user wishes to mine patterns that are interesting and specific to their application point of view. This experimental section can be divided into subsections, the contents of which vary according to the subject matter of the article. It must contain all the information about the experimental procedure and materials used to carry out experiments.

5. APPLICATION OF INCREMENTAL MINING OF FREQUENT PATTERNS

- **Business:** Pattern and trend analysis of customers. e.g.: Shopping patterns of customers which can help the retailer to know the need of customers and as a future action stores layout may change accordingly. The customer data keeps growing with time. This growing data can be mined incrementally using our research; Stock analysis by

investors. e.g.: Investors can find out stocks with good performance etc.

- **Education:** Behavior analysis of students both in offline and online environment; Identification of the risk level of students; Enrollment Management (branch of study, courses, and programs in which they can enrol) and Curriculum Development which can help planners design a curriculum focused on the demands of the future workforce etc. so that successful, goal-oriented and quality education can be obtained.
- **Banking and Finance:** Patterns and association analysis specially loan payment and credit analysis of customers; unusual pattern analysis etc.

6. CONCLUSION

To address the issues arising from real time increasing data, the research need to focus on incremental mining that helps in the decision-making process. This state of the art paper reviews methods of Incremental Mining of frequent patterns and highlights the possible advantages and limitation of each method. Incremental Mining of frequent patterns approaches discussed here is Apriori based Incremental Mining Algorithm, Partition based Incremental Mining Algorithms, Tree/Pattern Growth based Incremental Mining, High Utility Incremental Frequent Pattern Mining and Constraint based Incremental Mining. This is a state-of-art paper, where have discussed five Research Questions, Goal-related issues and Application of Incremental Mining of Frequent Patterns in real world scenario. This work will help researchers to find out related advances, limitations of each approach and suitable methods as per their research requiremetnts.

6. REFERENCES

[1] Lim, A.H.L, & Lee, C.S. (2010). Processing online analytics with classification and association rule mining. *Knowledge-Based Systems*, 23(3), 248-255, doi:10.1016/j. knosys.2010.01.006.

[2] Han, J., & kamber, M. (2006). Data Mining Concepts and Techniques. Morgan Kaufmann Publishers, An imprint of Elsevier, San Francisco, CA 94111.

[3] Han, J., Cheng, H., Xin, D., & Yan, X. (2007). Frequent pattern mining: current status and future directions. *Data Mining and Knowledge Discovery*, *15*(1), 55-86. doi:10.1007/s10618-006-0059-1.

[4] Su, M.Y., Yu, C.J., & Lin, C.Y. (2009). A real time network intrusion detection system for large-scale attacks based on an incremental mining approach. *Computers & Security, Science Direct*, *28*(5), 301-309. doi: 10.1016/j.cose.2008.12.001.

[5] Cameron, J.J., & Leung, C.K. (2011). Mining Frequent Patterns from Precise and Uncertain Data. *Journal of Systems and Computer*, *1*(1), 3-22.

[6] Cheung, W., Han, J., Ng, T., Wong, C.Y. (1996). Maintenance of discovering association rules in large databases: an incremental updating technique. In *proceedings of 12th IEEE International Conference on Data Engineering (*pp.106-114), New Orleans, L.A., February 26-March 01. doi: 10.1109/ICDE.1996.492094.

[7] Cheung,W., Lee S.D, & Kao B. (1997). A general incremental technique for maintaining discovered association rules. *In proceedings of 5th International Conference on Database Systems for Advanced Applications* (pp.185-194), Australia.

[8] Ayan, N.F., Tansal, A.U., Arkun, M.E.(1999). An efficient algorithm to update large itemsets with early pruning. In *proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp.287-291),USA, doi: 10.1145/312129.312252.

[9] Veloso, A., Pôssas, B., Meira, W., Márcio, Jr., Carvalho, B. (2001). Knowledge management in association rule mining, In *proceedings of Integrating Data Mining and Knowledge Management, held in conjunction with the 2001 IEEE International Conference on Data Mining*, doi: 10.1.1.63.9197.

[10] Lee, C.H., Lin, C.R., & Chen, M.S. (2001). Sliding-window filtering: An efficient algorithm for incremental mining. In *proceedings of the 10th International Conference on Information and Knowledge Management* (pp. 263-270), November 5-10, Atlanta, Georgia.

[11] Chang,C.H., & Yang S.H., (2003). Enhancing SWF for incremental Association Mining by itemset Maintenance. In *proceedings of the 7th Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 301-312), Seol, Koria, April 30- May 2, doi: 10.1.1.7.7783

[12] Cheung, W., & Zaïane, O.R. (2003). Incremental Mining of Frequent Patterns without Candidate Generation or Support Constraint. In *proceedings of IEEE 7th International Conference on Database Engineering and Applications Symposium* (pp. 111-116). Hong Kong, July 18. doi: 10.1109/IDEAS.2003.1214917.

[13] Leung, C.K.S., Khan, Q.I, Li, Z. & Hoque, T. (2007). CanTree: a canonical-order tree for incremental frequent- pattern mining. *Knowledge and Information Systems*, *11*(3), 287-311. doi:10.1007/s/0115-006-0032-8.

[14] Hong, T.P., Lin, C.W., & Wu, Y.L. (2008). Incrementally fast updated frequent pattern trees. *Expert Systems with Applications*, *34*, 2424-2435. doi: 10.1016 /j.eswa. 200.04.009.

[15] Yang, C.H., & Yang, D.L. (2009). IMBT-A Binary Tree for Efficient Support Counting of Incremental Data Mining. In *proceedings of International Conference on Computational Science and Engineering* (pp. 324-329). IEEE Computer Society Washington, DC, USA, August 29-31. doi: 10.1109/CSE.2009.360.

[16] Jian-ping,L., Ying,W., Fan-ding,Y.(2010).Incremental Mining Algorithm Pre-Fp in Association Rules based on FP-Tree. *in Proc. IEEE Int. Conf. On Networking and Distributed Computing,* Hangzhou, China , (pp. 199 – 203), DOI: 10.1109/ICNDC.2010.49

[17] Prajakta, V. & Sane, S. S.(2012).Incremental learning Algorithm for Association Rule Mining. *International Journal of Scientific & Engineering Research*, Vol.3, Issue 11, pp. 1-5.

[18] Jindal, R. & Dutta Borah, M. (2016). A novel approach for mining frequent patterns from incremental data. *International Journal of Data Mining, Modelling and Management*, Inderscience Publications, *8*(3), doi: 10.1504 /IJDMMM.2016. 079071.

[19] Yun, U., & Leggett, J.J. (2005). WFIM: Weighted Frequent Itemset Mining with a Weight Range and a Minimum Weight. In *proceedings of the SIAM International Conference on Data Mining* (pp. 636-640). Newport Beach, CA, April 21-23. doi: 10.1137/1.9781611972757.76.

[20] Yun, U., & Leggett, J.J. (2006). WIP: Mining Weighted Interesting Patterns with a Strong Weight and/or Support Affinity. In *proceedings of the SIAM International Conference on Data Mining* (pp. 623-627). Bethesda, April 20-22.

[21] Vo, B., Coenen, F., & Le, B. (2013). A new method for mining frequent weighted itemsets based on WIT-trees. *Expert Systems with Applications*, *40*(4), 1256-1264, doi:10.1016/j.eswa.2012.08.065.

[22] Ahmed, C.F., Tanbeer, S.K., Jeong, B.S., Lee, Y.K., & Choi, H.J. (2012). Single-pass incremental and interactive mining for weighted frequent patterns. *Expert Systems with Applications*, *39*(9), 7976-7994. doi: 10.1016/j.eswa.2012.01.117.

[23] Tseng, V.S., Shie, B.E, Wu, C.W., & Yu, P.S. (2013). Efficient Algorithms for Mining High Utility Itemsets from Transactional Databases. *IEEE Transactions on Knowledge and Data Engineering*, *25*(8), 1772-1786.

[24] Fournier-Viger, P., Wu, C., Zida, S., & Tseng, V. S. (2014). FHM: Faster high-utility itemset mining using estimated utility co-occurrence pruning. *In Foundations of Intelligent Systems. Lecture notes in computer science,* Springer International Publishing, *8502*, 83-92. doi: 10.1007/978-3-319-08326-1_9.

[25] Yun, U., Ryang, H., & Ryu, K.H. (2014). High Utility itemset mining with techniques for reducing overestimated utilities and pruning candidates. *Expert Systems with Applications*, *41*(8), 3861-3878. doi: 10.1016/j.eswa.2013.11.038.

[26] Patil, S. K., & Bojewar, S. (2017). Application of high utility mining for pattern prediction. 2017 International Conference on Trends in Electronics and Informatics (pp. 1154-1158). 11-12 May, Tirunelveli, India, doi: 10.1109/ICOEI.2017.8300895.

[27] Lin, J.C-W., Ren, S., Fournier-Viger, P., (2018) MEMU: More Efficient Algorithm to Mine High Average-Utility Patterns with Multiple Minimum Average-Utility Thresholds, *IEEE Access* (99), doi 10.1109/ACCESS.2018.2801261.

[28] Chiu, S.Y, Chiu, S.C. & Huang, J.L. (2009). On Mining Repeating Pattern with Gap Constraint. In *proceedings of 11th IEEE International Conference on High Performance Computing and Communications* (pp.557-562), Korea University, Seoul, Korea, June 25-27. doi:10.1109/HPCC.2009.65.

[29] Ruggieri, S. (2010). Frequent Regular Itemset Mining. In *proceedings of the 16th ACM International Conference on Knowledge Discovery and Data Mining* (pp. 263-272). Washington, DC, July 25-28.

[30] Bhatt, U.Y., & Patel, P.A. (2015). Mining Interesting Rare Items with Maximum Constraint Model Based on Tree Structure. In *proceedings of 2015 Fifth International Conference on Communication Systems and Network Technologies* (pp. 1065-1070). Gwalior, India, April 4-6. doi:10.1109/CSNT.2015.190.

[32] Dutta Borah, M. & Jindal, R. (2015). Constraint based filtering for Incremental Data Mining", In *proceedings of the 7th IEEE International Conference on Computational Intelligence and Communication Networks*, Jabalpur, India, December 12-14.

[33] Qiu, P., Zhao, L., & Dong, X. (2017) NegI-NSP: Negative sequential pattern mining based on loose constraints. , 43rd Annual Conference of the IEEE Industrial Electronics Society, 29 Oct.-1 Nov., Beijing, China, doi. 10.1109/IECON.2017.8216579.

[34] Wu,Y., Tong,Y. & Zhu,X. ,NOSEP (2018): Nonoverlapping Sequence Pattern Mining With Gap Constraints. IEEE Transactions on Cybernetics (IEEE Early Access article) (99), 1-14.doi. 10.1109/TCYB.2017.2750691.

How to cite this article:
Dutta Borah M, Jindal R. A study on Incremenatal Mining of Frequent Patterns. J. Fundam. Appl. Sci., 2018, *10(2), 97-112.*

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325168839

Assessment of Academic Departments Efficiency Using Data Envelopment Analysis

Article · May 2017

CITATIONS
0

READ
1

1 author:



Nand Kumar
Delhi Technological University
35 PUBLICATIONS **43** CITATIONS

SEE PROFILE

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

Efficiency analysis banks in India. View project

Dates Extended for Full manuscript Submission at ICAPIE 2016, Delhi, India View project

All content following this page was uploaded by Nand Kumar on 16 May 2018.

The user has requested enhancement of the downloaded file.

Available online at www.ijapie.org

INTERNATIONAL JOURNAL OF ADVANCED PRODUCTION AND INDUSTRIAL ENGINEERING

IJAPIE-2017-02-215, Vol 2 (2), 32-41

IJAPIE
Connecting Science & Technology with Management.
A Journal for all Products & Processes.

Assessment of Academic Departments Efficiency Using Data Envelopment Analysis

Sushil Yadav*, Lakshay, Chitransh Joshi, Mandeep Singh, Gurmeet Singh Bhatia , Ranganath M. Singari, Nand Kumar
(Delhi Technological University, Delhi, India)
**Email: sushilyadav5050@gmail.com*

Abstract : Data Envelopment Analysis (DEA) is a relatively new data-oriented approach for evaluating the technical efficiency of a set of peer entities called Decision Making Units (DMUs). A college administration benefits from measurement of relative efficiency of their departments for guiding and supporting inefficient departments. In this study, efficiency measurement of some academic departments within an University. Data Envelopment Analysis (DEA) approach is used for measurement. CCR (Charnes et al. 1978) model of Data Envelopment analysis has been applied for assessing the efficiency of different academic departments using the software "Frontier Analyst". Furthermore, after applying this model for assessment of academic departments efficiency, it was found that Biotechnology Department (Efficiency- 54.8%), Applied Physics Department (Efficiency-89.4%), Civil Department (Efficiency-93.8%), and ECE Department (Efficiency-99.4%) are inefficient Decision Making Units.

Keywords: Data Envelopment Analysis (DEA), technical efficiency, Decision Making Units (DMUs), relative efficiency, CCR model, Frontier Analyst,

1. INTRODUCTION

Higher education is the backbone of development and economic growth in any country. Given that the academic institutions are responsible for the capacity building required for a country's long-term plans, the educational system in particular, is one of the pillars a country depends on to increase its productivity and thus efficiently implement its strategic plans. There is no doubt that an efficient system of institutions of higher education is crucial to providing the necessary profession manpower of scientists, engineers, doctors, and teachers. Therefore, there is a need to assess the efficiency of the educational institutions, and whether the high cost being spent on them is worthwhile. Further, it is necessary to have standards by which all educational institutions could be questioned through the evaluation of efficiency of using resources (Inputs) and achieving the goals (Outputs) for which these resources were spent.

Delhi Technological University (DTU), formerly known as Delhi College of Engineering (DCE), is a premier government university located in New Delhi, India. It is one of the oldest engineering colleges in India and Delhi's first engineering college. It was established in 1941 as Delhi Polytechnic and was under the control of the Government of India. The seeds of Delhi Polytechnic were sowed by the Wood and Abott Committee of 1938. The technical school was created to cater to the demands of Indian industries. At that time, Delhi Polytechnic offered courses in Arts, Architecture, Commerce, Engineering, Applied Science and Textiles. The National Diploma awarded by Delhi Polytechnic was considered an equivalent to the B.E. Degree by the then UPSC. The college has been under the government of the National Capital Territory of Delhi since 1963 and was affiliated with the University of Delhi from 1952 to 2009. In 2009, the college was given state university status, thus changing its name to Delhi Technological University. Till the year 2009, DCE shared its admission procedure and syllabus for various B.E courses with there another branch known as Technology, formerly (DIT), of which were prescribed by Faculty of Technology, University of Delhi. It offers courses towards Bachelor of Technology (B.Tech), Master of Technology (M.Tech), and Doctor of Philosophy (Ph.D.) and contains ten academic departments offering fifteen undergraduate programs namely Computer Engineering, Electronics and Communication, Electrical and Electronics Engineering, Electrical Engineering, Information Technology, Mathematics and Computing Engineering, Software Engineering, Mechanical engineering, Automobile engineering, Civil Engineering, Production and Industrial Engineering, Engineering Physics, Polymer science and Chemical Technology, Biotechnology, Environmental Engineering with a strong emphasis on scientific and technological education and research.

Data Envelopment Analysis (DEA) is a relatively new data-oriented approach for evaluating the technical efficiency of a set of peer entities called Decision Making Units (DMUs). DEA provides a single measure and easily deals with multiple inputs and multiple outputs. Since the DEA technique was first developed, it has been widely applied to industries as diverse as health care (Bhat, Verma, & Reuben,

2001; Jacobs, Smith, & Street, 2006), Banking (Hassan & Sanchez, 2007), and transportation (Pathomsiri, 2006) and many other industries and organizations. Further, DEA-approach has proved especially valuable in cases where we have non-marketed inputs or outputs and/or cannot be derived or agreed upon among different DMUs.
The objective of this study is the "Assessment of Academic Departments Efficiency Using Data Envelopment Analysis"

2. LITERATURE REVIEW

2.1 DATA ENVELOPMENT ANALYSIS

When analyzing system performance, qualitative evaluations such as ``good", ``fair", ``adequate", and ``poor" are vague and difficult to utilize in any meaningful way. As a result, quantitative performance measures are often preferred to such qualitative evaluations. A numerical performance measure might be utilized because the data are readily available, or because it has been used for a long time. Data Envelopment Analysis is one among the several mathematical models to measure flexibility.

DEA involves an alternative principle for extracting information about a population of observations such as those shown in Figure 1. In contrast to parametric approaches whose object is to optimize a single regression plane through the data, DEA optimizes on each individual observation with an objective of calculating a discrete piecewise frontier determined by the set of Pareto-efficient DMUs. Both the parametric and non-parametric (mathematical-programming) approaches use all the information contained in the data. In parametric analysis, the single optimized regression equation is assumed to apply to each DMU. DEA, in contrast, optimizes the performance measure of each DMU. This results in a revealed understanding about each DMU instead of the mythical "average" DMU. In other words, the focus of DEA in on the individual observations as represented by the n optimizations (one for each observation) required in DEA analysis, in contrast to the focus on the averages and estimation of parameters that are associated with single-optimization statistical approaches.

The parametric approach requires the imposition of a specific functional form (e.g., a regression equation, a production function, etc.) relating the independent variables to the dependent variables(s). The functional form selected also requires specific assumptions about the distribution of error terms (e.g., independently and identically normally distributed) and many other restrictions, such as factors earning the value of their marginal product. In contrast, DEA does not require any assumption about the functional form. DEA calculates a maximal performance measure for each DMU relative to all other DMUs in the observed population with the sole requirement that each DMU lie on or below the extremal frontier. Each DMU not on the frontier is scaled against a convex combination of the DMUs on the frontier facet closest to it.

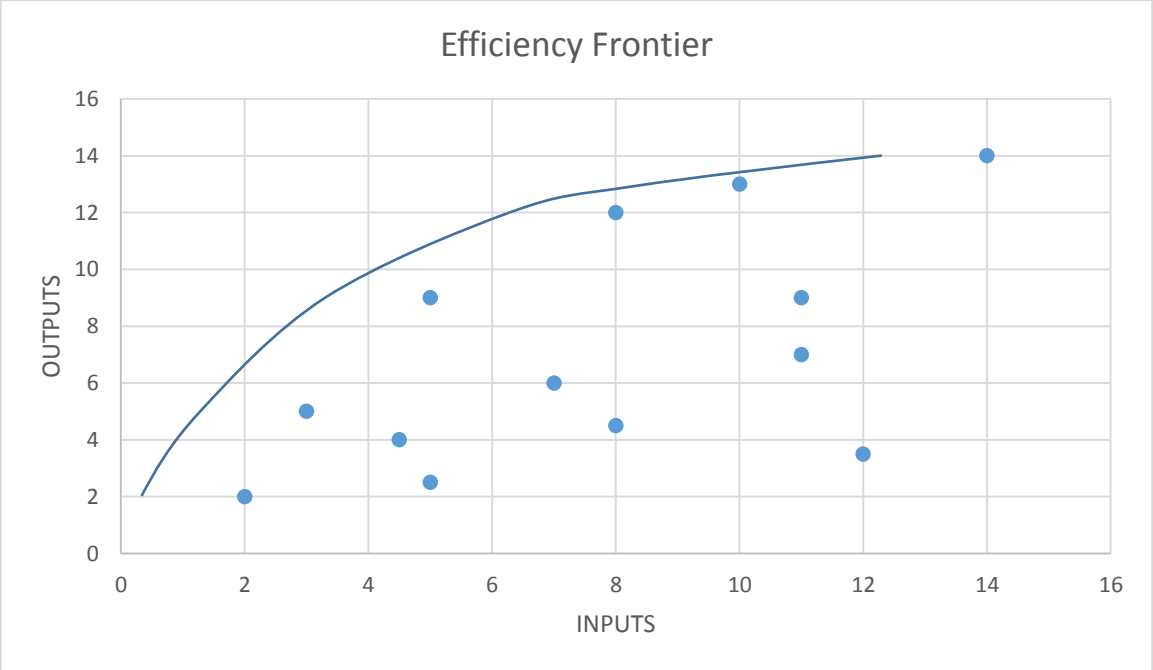


Figure 1

Charnes, Cooper, and Rhodes (1978) extended Farrell's (1957) idea linking the estimation of technical efficiency and production frontiers. Their CCR model generalized the single-output/input ration measure of efficiency for a single DMU in terms of a fractional linear-programming formulation transforming the multiple input/output characterization of each DMU to that of a single "virtual" output and virtual input. The relative technical efficiency of any DMU is calculated by forming the ratio of a weighted sum of outputs to a weighted sum of inputs, where the

weights (multipliers) for both outputs and inputs are to be selected in a manner that calculates the Pareto efficiency measure of each DMU subject to the constraint that no DMU can have relative efficiency score greater than unity.

The solid line in Fig.-1 represents a frontier derived by DEA from data on a population of DMUs, each utilizing different amounts of a single input to produce various amounts of a single output. It is important to know that DEA calculations, because they are generated from actual observed data for each DMU, produce only relative efficiency measures. The relative efficiency of each DMU is calculated in relation to all the other DMUs, using the actual observed values for the outputs and inputs of each DMU. The DEA calculations are designed to maximize the relative efficiency score of each DMU, subject to the condition that the set of weights obtained in this manner for each DMU must also be feasible for all the other DMUs included in the calculation. DEA produces a piecewise empirical extremal production surface (e.g., the solid line in figure 1), which in economic terms represent the revealed best practice production frontier – the maximum output empirically obtainable from any DMU in the observed population, givens its level of inputs.

For each inefficient DMU (one that lies below the frontier), DEA identifies the sources and level of inefficiency for each of the inputs and outputs. The level of inefficiency is determined by comparison to a single referent DMU or a convex combination of other referent DMUs located on the efficiency frontier that utilize the same level of inputs and produce the same or higher level of outputs. This is achieved by requiring solutions to satisfy inequality constraints that can increase some outputs (or decrease some inputs) without worsening the other inputs or outputs. The calculation of potential improvement for each inefficient DMU does not necessarily correspond to the observed performance of any actual DMU making up the piece wise production frontier or to a deterministic projection of an inefficient DMU onto the efficient frontier. The calculated improvements (in each of the inputs and outputs) for inefficient DMUs are indicative of potential improvements obtainable because the projections are based on the revealed best practice performance of "comparable" DMUs that are located on the efficient frontier.

To conclude this section by noting that, initially, DEA relative efficiency solutions were of interest to operational analysts, management scientists, and industrial engineers largely because of three features of this method:

1. Characterization of each DMU by a single summary relative efficiency score;
2. DMU-specific projections for improvements based on observable referent revealed best-practice DMUs; and
3. Obviation by DEA of the alternative and indirect approach of specifying abstract statistical models and making inferences based on residual and parameter coefficient analysis.

The attraction of DEA to traditional Frontier econometricians (Schmidt, 1986, for early concerns) emerged from the new insights obtained in production frontier analysis involving their existence and the variance around them. For example, the BCC model relaxed the constant-returns-to-scale requirement of the original CCR ratio model and made it possible to investigate local returns to scale. Thus it became evident that DEA can complement information about average returns to scale from econometric models with DMU-specific information for most productive scale size (Banker, 1984) and allocative efficiency (Banker and Mainderata, 1988) in contrast to only average most-productive scale and average allocative efficiency obtained from traditional econometric frontier analysis.

2.2 DECISION MAKING UNITS "DMUs"

Data Envelopment Analysis, a non-parametric performance measurement technique, is used for determining the relative efficiency of a set of comparable business called Decision Making Units (DMUs) and it can also be used for benchmarking.

The Delhi Technological University (DTU) has ten departments offering fifteen programs. The ten departments namely Applied Chemistry, Applied Mathematics, Applied Physics, Biotechnology, Civil Engineering, Computer Science & Engineering, Electrical Engineering, Mechanical Engineering, Electronics and Communication Engineering, and Environment Engineering are considered as Decision Making Units (DMUs).

Hence, the research sample includes 10 DMUs offering 15 programs as shown in Table 1.

Table 1 – Decision Making Units

S. No.	Decision Making Units (DMUs)
1.	Applied Chemistry
2.	Applied Mathematics
3.	Applied Physics
4.	Biotechnology
5.	Civil Engineering
6.	Computer Science & Engineering
7.	Electrical Engineering
8.	Mechanical Engineering
9	Electronics and Communication Engineering
10	Environment Engineering

2.4 VARIABLES SELECTION

The efficiency of organizations has been studied by many researchers in different contexts, including research units and university departments. What make efficiency measurement in universities an important issue are legislative requirements, difficulty in determining input and output measures, large number of such measures relative to number of departments, and data collection problems. All these become even more difficult when a college administration desires to compare efficiencies of different academic departments within the college.

A college administration benefits from measurement of relative efficiency of their departments for guiding and supporting inefficient departments. In this study, efficiency measurement of some academic departments within Delhi Technological University. Data Envelopment Analysis (DEA) approach is used for measurement. Necessary criteria and indicators to measure the efficiency of the engineering college departments are described.

Input and output variables: several input-output factors have been considered for this comparative study, they are being enlisted below along with their respective importance and specific roles so as to better appreciate their impact on the overall working efficiency of an academic department.

1. Total Number of Faculty Members (Input 1) - total number of faculty members in any department is one of the most rudimentary inputs by the University in improving the efficiency of the outputs produced. This input has been chosen in the place of staff salaries as suggested by GU¨LSER KO¨KSAL and BURAK NALC¸ACI whose study on Turkish Universities has served as the guiding document. To present a wholesome perspective weighted scores have been given(professor=3pts, associate professor=2pts, assistant prof=1pts)

2. Infrastructure (Input 2) – in order to inculcate the available infrastructure we have included the number of labs that a department has, although several studies have included the area component (Kao and Hung,2006) into their assessment due to empirical constraints we chose the number of labs as the input.

3. Operational Budget (Input 3) – operational budget allotted to different departments can act as a key factor in affecting the output (Kao and Hung,2006) , and is thus included as our last input for this study. Net financial allocations have been directly linked to academic outputs.

4. Total Number of Admissions (Output 1) – total number of admissions is one of the most suited output that an academic department may generate, maximizing it with a given set of inputs can improve efficiency without incurring extra cost. Weighted scores have been included here (Undergraduate= 1pts, Post Graduate= 2pts, PhD candidate= 3pts).

5. Research publications (Output 2) – measuring the efficiency of any academic department shall have research publications as its focus areas and thus this is a desired output.

6. Public Service Activities (Output 3) - the final output is Public Service Activities (PSA)which includes the number of officially documented workshops, conferences, training courses and other extracurricular activities by the teaching staff of each department, that are officially documented at the university (Salah R. Agha, Ibrahim Kuhail ,2011).

2.6 Basic Efficiency Concepts

Efficiency can be simply defined as the ratio of output to input. More output per unit of input reflects relatively greater efficiency. If the greatest possible output per unit of input is achieved, a state of absolute or optimum efficiency has been achieved and it is not possible to become more efficient without new technology or other changes in the production process.

The most common efficiency concept is

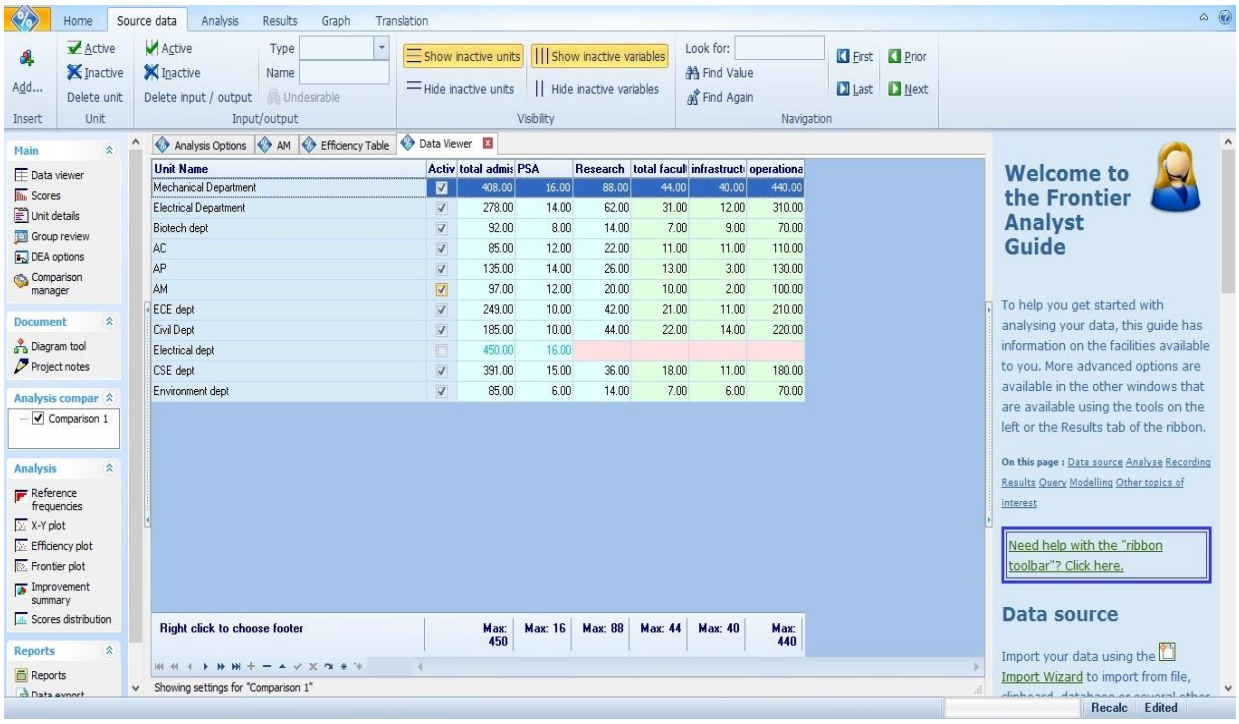
Technical Efficiency: the conversion of physical inputs (such as the services of employees and machines) into outputs relative to best practice. In other words, given current technology, there is no wastage of inputs whatsoever in producing the given quantity of output. An organization operating at best practice is said to be 100% technically efficient. If operating below best practice levels, then the organization's technical efficiency is expressed as a percentage of best practice. Managerial practices and the scale or size of operations affect technical efficiency, which is based on engineering relationships but not on prices and costs.

Allocative Efficiency refers to whether inputs, for a given level of output and set of input prices, are chosen to minimize the cost of production, assuming that the organization being examined is already fully technically efficient. Allocative efficiency is also expressed as a percentage score, with a score of 100% indicating that the organization is using its inputs in the proportions that would minimize costs. An organization that is operating at best practice in engineering terms could still be allocatively inefficient because it is not using inputs in the proportions which minimize its costs, given relative input prices.

Finally, **Cost Efficiency** refers to the combination of technical and allocative efficiency. An organization will only be cost efficient if it is both technically and allocatively efficient. Cost efficiency is calculated as the product of the technical and allocative efficiency scores (expressed as a percentage), so an organization can only achieve a 100%

score in cost efficiency if it has achieved 100% in both technical and allocative efficiency.

Table 2 – Data Collection



2.7 DEA MODEL

We can measure the efficiency of a decision making unit (DMU) simply by comparing its outputs to its inputs. The complex nature of the relation between multiple inputs and multiple outputs involved in the efficiency analysis of DMUs requires utilization of techniques which can handle large number of variables and constraints. For this purpose, Charnes et al. (1978) have proposed DEA. DEA basically discriminates efficient DMUs from inefficient ones. DEA can also point to the improvements each DMU needs to make in each input and output measure. Several studies compare DEA with other methods, and show its importance in measuring relative efficiency of DMUs over regression analysis (Cubbin and Tzanidakis 1998; Seiford and Thrall 1990; Thanassoulis 1993), ratio analysis (Thanassoulis et al. 1996), principle component analysis (Zhu 1998), cluster analysis and discriminant analysis (Stern et al. 1994).

Ali et al. (1995) provide a unified framework for classification of the basic DEA models. DEA models are grouped into CCR (Charnes et al. 1978), BCC (Banker et al. 1984) and Additive Models (Charnes et al. 1985). A DEA model basically identifies among all DMUs the ones that determine an envelopment surface or efficient frontier. The models are classified into two based on the type of envelopment surface: constant returns to scale (such as CCR model) and variable returns to scale (such as BCC model). The CCR model is appropriate when the efficiency of a DMU is not affected by its size (number of personnel, students, and so on). The BCC model, on the other hand, accommodates scale effects in the analysis. DEA models are also grouped into two, depending on the type of proportional movement towards the efficient frontier: input oriented models and output oriented models. Input oriented models address the question "By how much can input quantities be proportionally reduced without changing the output quantities produced?" whereas output oriented models address the question "By how much can output quantities be proportionally expanded without altering the input quantities used?"

The aim of this study is to develop a system to measure the efficiency of college departments and guide the inefficient ones by showing how much they should improve their outputs by using the same amount of inputs or how much they should reduce their inputs by producing the same amount of outputs. Thus, both the input oriented and output oriented model are considered in this case for assessment of the efficiencies of academic departments. Furthermore, the CCR model is preferred, due to less computational effort required by it, and the fact that the number of the departments is very small. The efficiency measure of the output oriented CCR model reflects the ability of a department to obtain maximum output from a given set of inputs and the efficiency measurer of the input oriented CCR model reflects the ability of a department to obtain a given output from minimum set of inputs.

DEA is used to find the relative efficiency of a set of comparable businesses called Decision Making Units (DMU). In DEA efficiency is defined as:

$$\text{Efficiency} = \frac{weighted\ sum\ of\ output}{weighted\ sum\ of\ input}$$

The weight attached to each input or output is not decided beforehand, instead they are chosen to each unit under comparison in its most favorable form. The envelope, or frontier, becomes the surface linking all units whose relative efficiency cannot be exceeded. By definition those units are thus given 100% efficiency, the best possible efficiency for other units in the sample then bring them as close as possible to the envelope. Thus DEA tells us the efficiency of any given DMU relative to best performing units in the sample based on the inputs and outputs.

If there are 'n' DMUs, each with 'm' inputs and 's' outputs, the relative efficiency score of a test DMU 'p' is obtained as follows:

DMU 'p' is obtained as follows:

$$= \text{MAX.} \frac{\sum_{k=1}^{s} V(k).Y(kp)}{\sum_{j=1}^{m} U(j).X(jp)}$$

Such that $\frac{\sum_{k=1}^{s} V(k).Y(ki)}{\sum_{j=1}^{m} U(j).X(ji)} \leq 1$ for all i

(Ref Research Paper 2)

$V(k), U(j) \geq 0$ for all j,k

Where:
K= 1 to's'; j=1 to 'm'; i=1 to 'n'
Y(ki) = amount of output 'k' by DMU 'i'
X(ji) = amount of input 'j' used by DMU 'i'
V(k) = weight given to output 'k'
U(j) = weight given to input 'j'

This can be converted into a linear programming problem for ease of solving, thus, we have

$$\text{MAX} \left(\sum_{k=1}^{s} V(k).Y(ki) \right)$$

Such that $\sum_{j=1}^{m} U(j).X(ji) = 1$

And $\left(\sum_{k=1}^{s} V(k).Y(ki) - \sum_{j=1}^{m} U(j).X(ji) \right) \leq 0$

for all i (Ref Research Paper 2)

The above equations are run 'n' times (one for each DMU) to get the relative efficiency scores of the DMUs. A DMU is said to be efficient if the relative efficiency score is 1, a score <1 insures that it is inefficient. Each DMU selects input and output weights that maximize its relative efficiency.

2.7.1 Benchmarking in DEA

For every efficient DMU, DEA gives a set of efficient DMUs that can be used as benchmark for identifying areas of improvement. The benchmark can be obtained as follows:

Min E sub to –

$$\sum_{i=1}^{n} Q(i)\ Y(ki) \geq Y(kp)$$ for all j

$$\sum_{i=1}^{n} Q(i).X(ki) \leq E.\ X(kp)$$ for all k

$Q(i) \geq 0$ for all i (Ref Research Paper 2)

Where
E= efficiency score
Q(i) = Dual variable

These dual variables can be used to construct Hypothetical Composite Unit (HCU). HCU can be used to measure excess use of inputs or potential increase in outputs. If E= 1, no proof of inefficiency but if E<1 there is efficiency, and thus we can identify the areas of improvement in output and input sides.

DEA is receiving increasing importance as a tool for evaluating and improving the performance of manufacturing and service operations. It has been extensively applied in performance of manufacturing and service operations. DEA approach has the following benefits which make it suitable for its application in flexibility performance measurement and benchmarking.

1. DEA deals with individual cases (Madu and Kuei 1998).
2. It can produces a single measure for each company (Madu and Kuei 1998).
3. It places no restriction on the functional form of the input-output relationship.
4. Able to handle disproportionate multiple inputs and outputs (George and Rangaraj 2008).
5. Does not requiring the decision maker any priory arbitrary weights (George and Rangaraj 2008).
6. It focuses on revealed best-practice frontiers rather than on central tendency properties of empirical data (Madu and Kuei 1998)
7. It can provide an indication of the levels of improvement needed before an inefficient company could be considered efficient (Talluri 2000)

3. ASSESSMENT

3.1 EQUATIONS

Using these Input and output variables equations are formed for ten academic departments, identifying them as DMUs. The standard model of DEA is applied on the following set of equations to compute efficiency.

Testing for inefficiencies in DMU1 by checking for a dummy efficient DMU

Min. E, such that

$W_1 + W_2 + W_3 + W_4 = 1$

$O_{11}xW_1 + O_{12}xW_2 + O_{13}xW_3 + O_{14}xW_4 + O_{15}xW_5 + O_{16}xW_6 + O_{17}xW_7 + O_{18}xW_8 + O_{19}xW_9 + O_{110}xW_{10} >= O_{11}$
$O_{21}xW_1 + O_{22}xW_2 + O_{23}xW_3 + O_{24}xW_4 + O_{25}xW_5 + O_{26}xW_6 + O_{27}xW_7 + O_{28}xW_8 + O_{29}xW_9 + O_{210}xW_{10} >= O_{21}$
$O_{31}xW_1 + O_{32}xW_2 + O_{33}xW_3 + O_{34}xW_4 + O_{35}xW_5 + O_{36}xW_6 + O_{37}xW_7 + O_{38}xW_8 + O_{39}xW_9 + O_{310}xW_{10} >= O_{31}$

$I_{11}xW_1 + I_{12}xW_2 + I_{13}xW_3 + I_{14}xW_4 + I_{15}xW_5 + I_{16}xW_6 + I_{17}xW_7 + I_{18}xW_8 + I_{19}xW_9 + I_{110}xW_{10} <= I_{11}$ (E)
$I_{21}xW_1 + I_{22}xW_2 + I_{23}xW_3 + I_{24}xW_4 + I_{25}xW_5 + I_{26}xW_6 + I_{27}xW_7 + I_{28}xW_8 + I_{29}xW_9 + I_{210}xW_{10} <= I_{21}$ (E)
$I_{31}xW_1 + I_{32}xW_2 + I_{33}xW_3 + I_{34}xW_4 + I_{35}xW_5 + I_{36}xW_6 + I_{37}xW_7 + I_{38}xW_8 + I_{39}xW_9 + I_{310}xW_{10} <= I_{31}$ (E)

Testing for inefficiencies in other DMUs on similar lines we get the following results, as given below. Assessing them gives us the scope of improvement and identifies efficient DMUs, this helps in adopting techniques of operations management and productivity enhancement from the efficient DMUs to the relatively inefficient ones.

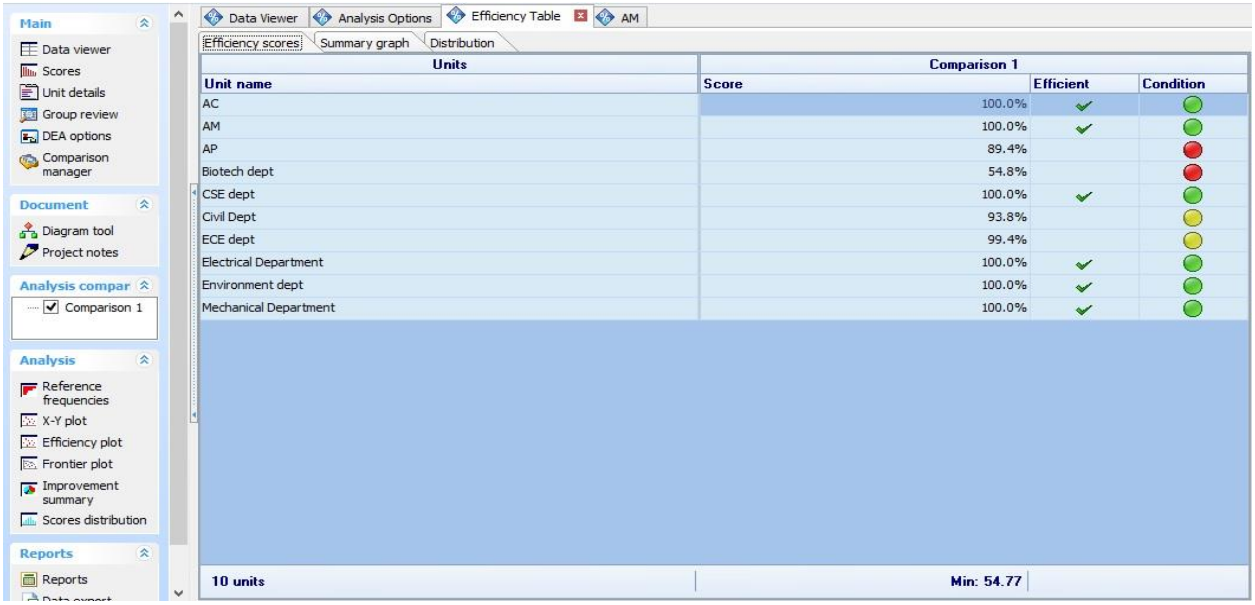


Figure 2 – Efficiency Score

As can be seen that the relative efficiencies have been color coded to identify the ones that need focus immediately, wherein red color denotes urgent attention with vast scope of improvement, yellow indicates a mild scope of improvement, and green indicates that the DMU lies on the efficiency frontier.

3.2 EFFICIENCY PLOTS

As Biotech dept. is the one with the least score it is analyzed first. Here it can be suggested that with an input optimization perspective either a reduction in the inputs is needed or a corresponding increase in the outputs is required to justify existing inputs.

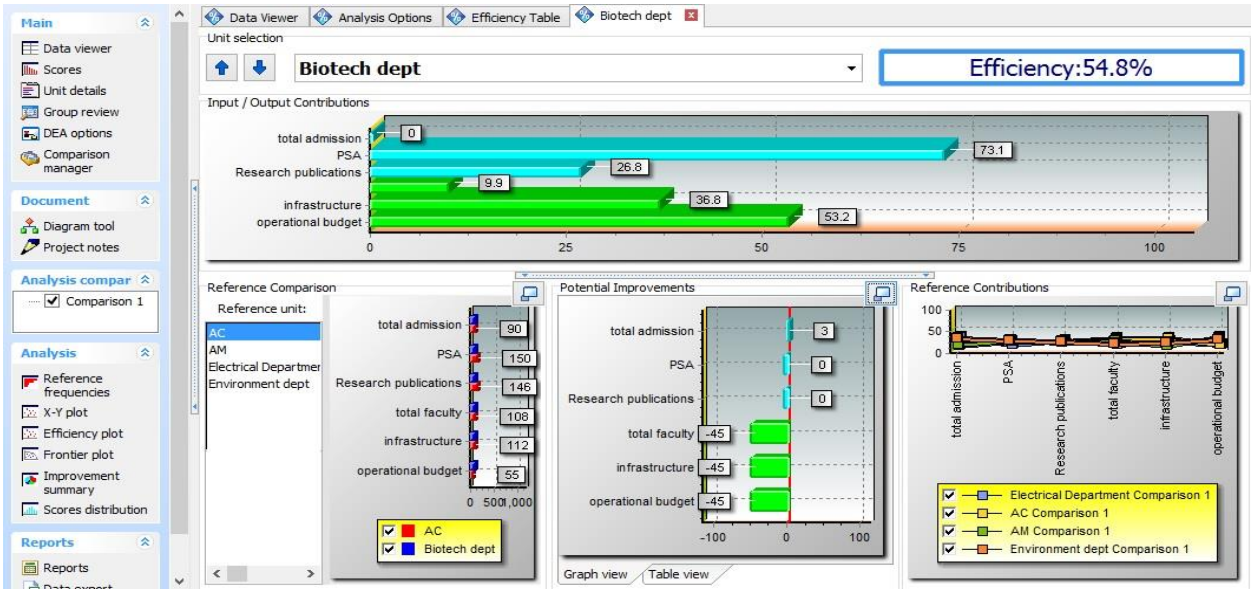


Figure 3 – Efficiency Plot of Biotechnology Department

Now looking at the efficiency record of Applied Physics department we can suggest a lay of in terms of faculty as the output levels don't justify the high levels of existing faculty. Suitable reductions in strength of the faculty and operational budget have been shown here.

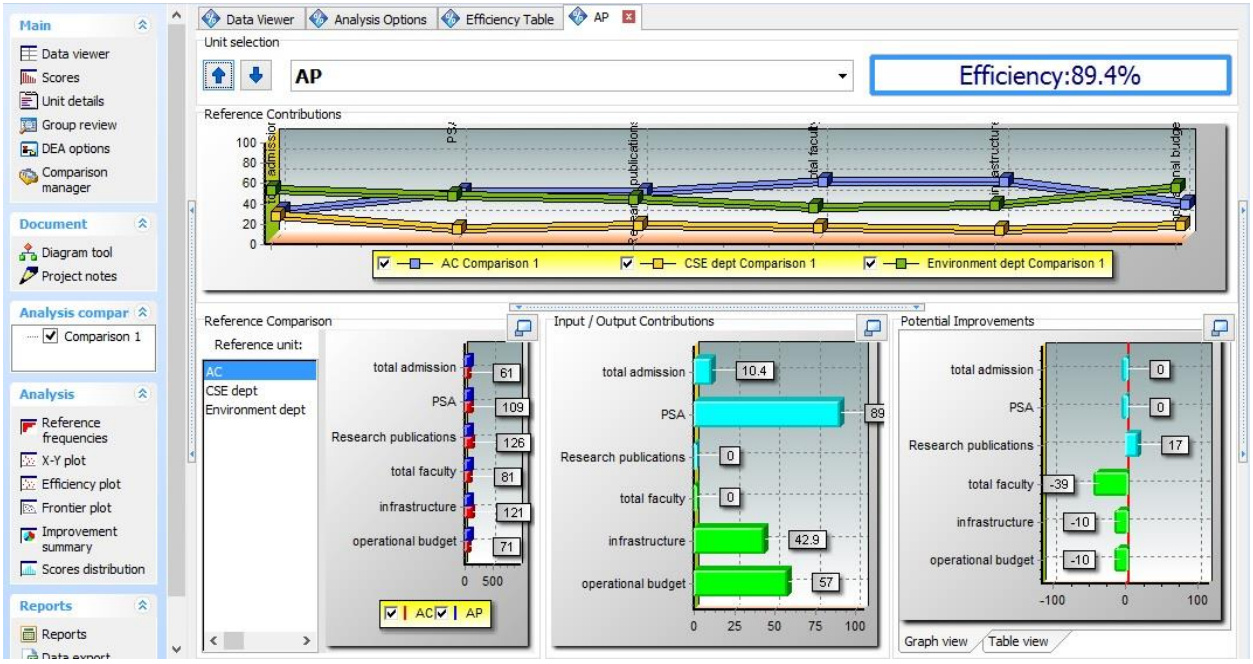


Figure 4 – Efficiency Plot of Applied Physics Department

Further taking an example from the yellow coded DMUs, we look at the Civil Department, here the efficiency relative to the others is fairly high (93.8%), but still a revision of the input given can be considered given the slight fall in efficiency.

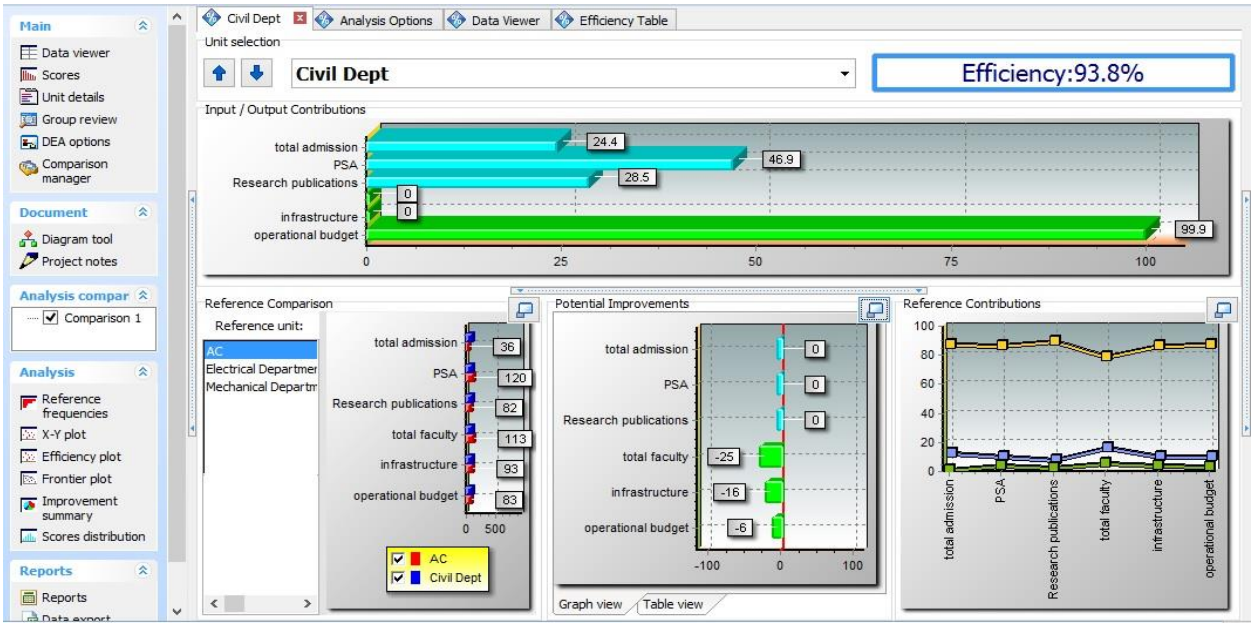


Figure 5 – Efficiency Plot of Civil Department

The total potential improvements for the Biotech Department have been specifically highlighted here in order to lay special emphasis on it due to its score. The best direction to start with in improvements would be to reduce the faculty and operational budget.

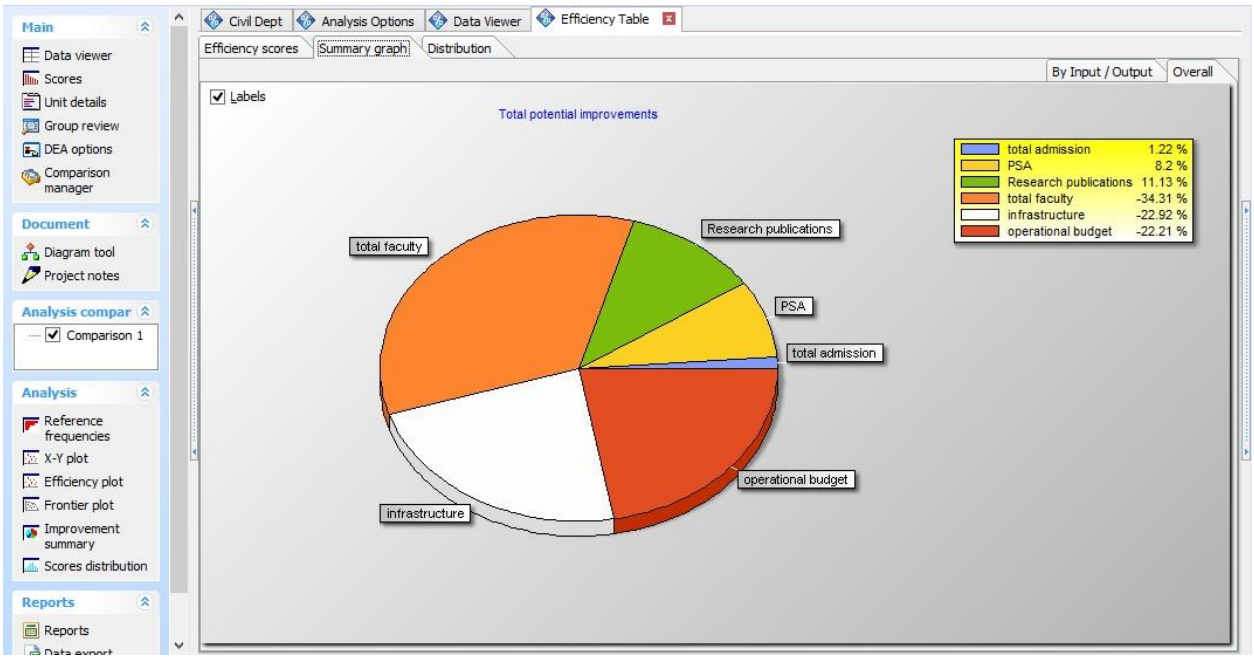


Figure 6 – Total Potential Improvements

4. CONCLUSIONS

During this project, we have gone through the research literature about Data Envelopment Analysis. CCR (Charnes et al. 1978) model of Data Envelopment analysis has been applied for assessing the efficiency of different academic departments using the software "Frontier Analyst". Furthermore, after applying this model for assessment of academic departments efficiency, we found that Biotechnology Department (Efficiency- 54.8%), Applied Physics Department (Efficiency-89.4%), Civil Department (Efficiency-93.8%), and ECE Department (Efficiency-99.4%) are inefficient Decision Making Units. Various steps should be taken to identify the bottlenecks in these cases and suggestion to rectify the same and thus improve the overall efficiency.

REFERENCES

1. Ying-Ming Wang, Kwai-Sang Chi: A neutral DEA model for cross-efficiency evaluation and its extension, in Expert Systems with Applications, An International Journal (2010),ref 3666–3675
2. Salah R. Agha, Ibrahim Kuhail, Nader Abdelnabi, Mahmoud Salem, Ahmed Ghanim: Assessment of academic departments efficiency using data envelopment analysis. Journal of Industrial Engineering and Management (2011), p301-325.
3. GU¨LSER KO¨KSAL & BURAK NALC¸ACI," The relative efficiency of departments at a Turkish engineering college: A data envelopment analysis ", International Journal of Operations & Production Management (2006) 51: 173–189
4. Rita Markovits-Somogyi," ranking efficient and inefficient decision making Units in data envelopment analysis ", Industrial Management & Data Systems, (2011)pp. 446 – 456
5. Abbott, M., & Doucouliagos, C. The Efficiency of Australian Universities: A Data Envelopment Analysis. Economics of Education Review, 22(1),(2003), 89-97..
6. Lopes, A., Lanzer, E. Data Envelopment Analysis-DEA And Fuzzy Sets to Assess The Performance of Academic Departments: A Study at Federal University Of Santa Catrina-UFSC. Pesquisa Operacional, 22(2), (2002), 217-230.
7. Moreno, A. & Tadepalli, R.. Assessing Academic Department Efficiency a Public University. Managerial and Decision Economics, 23(7), 385-397.8. Janusz Bucki Yvon Pesqueux, (2000),"Flexible workshop: about the concept of flexibility", International Journal of Agile Management Systems, Vol. 2 Iss 1, (2002) pp. 62 – 70
9. Tom Davis (1993) ,"Effective Supply Chain Management",Sloan Management Review, Summer 1993 edition.
10. Nand Kumar, Ranganath.M.S, Chandrashekhar, "Sustainability Analysis of Banking Sector in India using DEA", in National Conference on Inclusive & Sustainable Growth - IMT Nagpur
11. www.emeraldinsight.com
12. Banxia Frontier Analyst 4 Demo version.

View publication stats

Measurement 126 (2018) 134–142

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Boosting performance of power quality event identification with KL Divergence measure and standard deviation

Rajiv Kapoor[a], Rashmi Gupta[b], Le Hoang Son[c,d,*], Sudan Jha[e], Raghvendra Kumar[f]

[a] ECE Department, Delhi Technological University, India
[b] ECE Department, AIACT&R, GGSIPU, Delhi, India
[c] Institute of Research and Development, Duy Tan University, Danang, Viet Nam
[d] VNU University of Science, Vietnam National University, Viet Nam
[e] School of Computer Engineering, Kalinga Institute of Industrial Technology, Odisha, India
[f] Computer Science and Engineering Department, LNCT College, MP, India

ARTICLE INFO

Keywords:
Power quality
Event detection
KL Divergence
SVM
Standard deviation

ABSTRACT

Power quality event identification is widely recognized as one of the most interesting problems in electric engineering. It consists of two sub-problems: detection and classification. In the first step, a recognition algorithm is used to detect disturbance from power quality events. The next step classifies them into some groups by a machine learning method. In order to enhance the accuracy, a detection technique is required for classifying events in timely manner. In this paper, KL Divergence and Standard deviation are used within Support Vector Machine to detect and classify events. Experimental results with 12 events suggest a specific order of harmonic present in each event. KL Divergence and Standard deviation are obtained for voltage sag of 500 values and for harmonics with swell of 800 values. After calculating KL Divergence and Standard deviation, events are detected with more accuracy. The comparison shows that the new method achieves 94.02% of accuracy which is better than 92.33% of Abdoos et al. (2016), 93.47% of Ma et al. (2017), 89.92% of Li et al. (2016) and 93.87% of Kapoor et al. (2018).

1. Introduction

Power quality, which is also termed as electric power quality, involves voltage, frequency, and waveform [39]. It is defined as good power quality if supplied voltage remains steady within a specified range, and exhibits smooth voltage curve waveform [39]. It is vital to capture characteristics of events like sag, swell, etc. by monitoring power quality [18]. The power quality (PQ) event detection and classification (or identification) is widely recognized as one of the most interesting problems in electric engineering [6,16,51,1].

It consists of two sub-problems: detection and classification. In the first step, a recognition algorithm is used to detect disturbance from power quality events. Next, after successful detection of disturbances, the second step classifies them into some groups by a machine learning method like support vector machine (SVM) [17], Artificial Neural Network (ANN) [38], fuzzy logic [1], etc. It is recalled that good features are necessary for both power quality event detection and classification [29].

Regarding the first step, various methods have been used to detect power quality events such as Fourier transforms (FT) [49]. The power quality disturbances are non-stationary/non-periodic in nature so that they cannot be processed using Fourier transform [40]. The limitation is mainly due to quick variations (transient phenomenon) in incoming signals, which is difficult to observe visually by this method [40]. Short-Time Fourier Transform (STFT) was used for detecting these patterns [49]. This gives information in time domain and frequency domain with two window functions being used to obtain good resolution and increase operations [49]. The major drawback of this method is a fixed short window so that alternative approaches of wavelet change were utilized [9,14,34,40,38].

However, multi resolutions are decayed with the assistance of Low Pass and High Pass channels [17]. Wavelet has few weaknesses, for example, its confounded calculation to commotion level and reliance of exactness on picked wavelet [17]. This prompts the improvement of ST (Stockwell Change) [17] which is an invertible time recurrence unearthly limitation method joining the components of WT and STFT. ST utilizes an investigation window so that thickness is diminished with recurrence; thus giving a recurrence subordinate determination.

* Corresponding author at: 334 Nguyen Trai, Thanh Xuan, Hanoi, Viet Nam. Tel.: +84 904 171 284.
E-mail addresses: rajivkapoor@dce.ac.in (R. Kapoor), rashmig71@yahoo.com (R. Gupta), sonlh@vnu.edu.vn (L.H. Son), jhasudan@hotmail.com (S. Jha), raghvendraagrawal7@gmail.com (R. Kumar).

https://doi.org/10.1016/j.measurement.2018.05.053
Received 20 February 2018; Received in revised form 10 May 2018; Accepted 13 May 2018
0263-2241/ © 2018 Elsevier Ltd. All rights reserved.

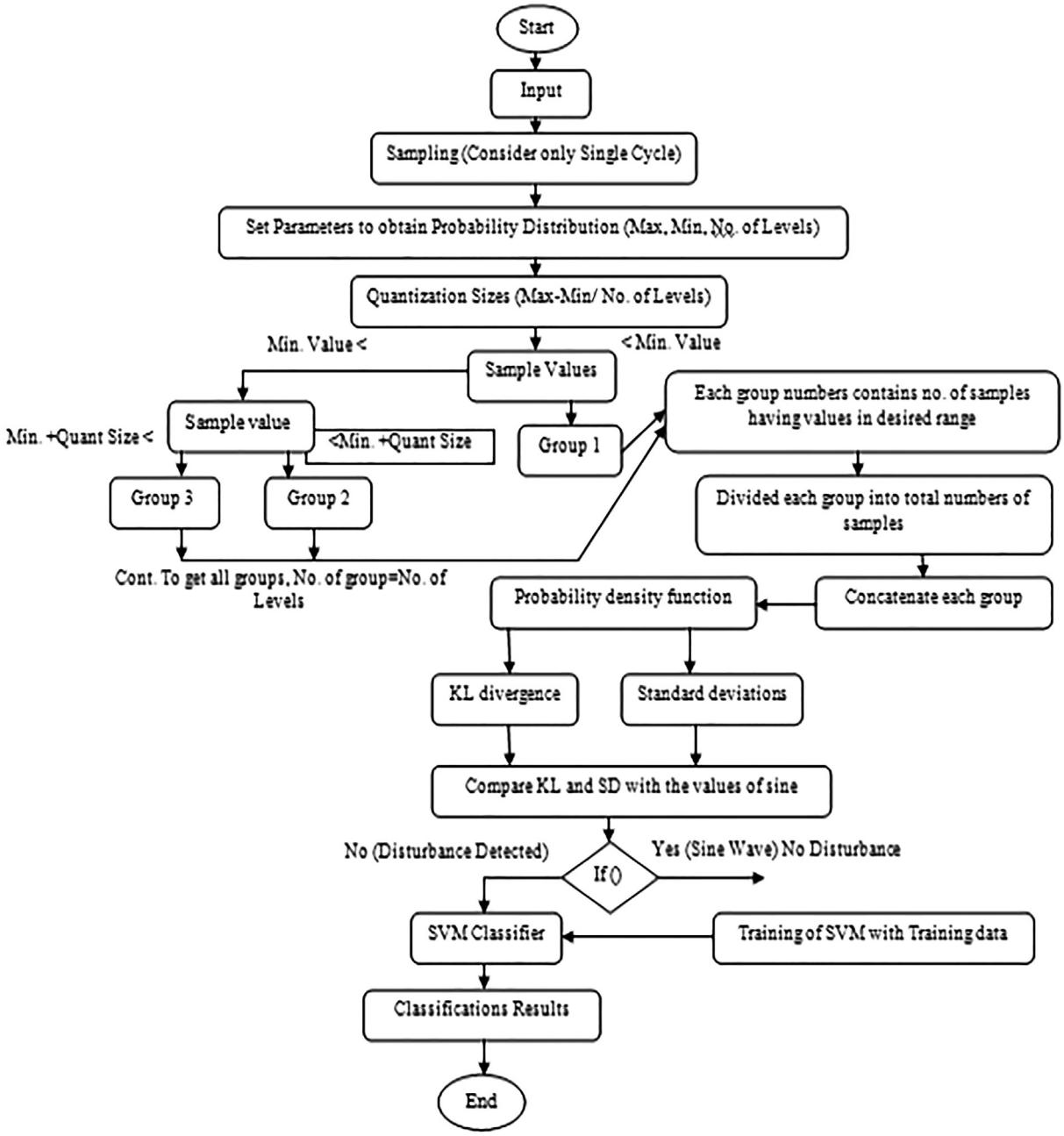


Fig. 1. Flowchart of the proposed method.

More literature review on power event detection can be referred in Section 2. However, we observe that time domain and frequency behaviors are two main approaches from those works [15]. In this paper, KL Divergence and Standard Deviation are used to detect and classify events quickly. KL Divergence acquires less processing time and can be applied to single cycle of the input signal instead of multiple cycles. It reduces the complexity and inefficient time. Standard Deviation is often used for input feature exaction from distorted voltage waveforms. Due to these reasons, KL Divergence and Standard Deviation methods have been more result-oriented approach.

In what follows, related studies and the proposed method are shown in Sections 2 and 3. Section 4 describes experiments. Section 5 delineates conclusions and further works.

2. Literature survey

Gupta and Kumar [10] proposed a straightforward and novel technique in view of symbolic dynamics for discovery of PQ aggravations. Hajian and Foroud [12] introduced a new feature selection method namely Orthogonal Forward Selection by incorporating Gram Schmidt and forward selection for detection of Power Quality Disturbances (PQDs). Kumar et al. [21] introduced the Hilbert – Huang transform for perceiving multiple events. Mohanty et al. [30] examined islanding discovery in a hybrid distributed generation framework by utilization of hyperbolic S-transform, time–time change and morphology techniques. Yong et al. [50] proposed a technique in view of a mix of binary classifiers which are advanced for those extraordinary situations where the genuine signs contain a large number of occasions inside the analyzed temporal window. Khokhar et al. [19] introduced an exhaustive survey on utilization of signal and optimization methods

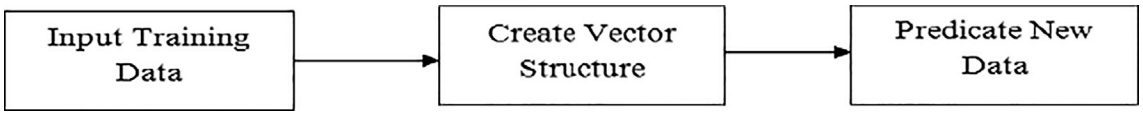


Fig. 2. Operating procedures of the classifier.

Table 1
List of power quality events.

PQ Disturbance	Model	Parameter
Sine	$f(t) = A\sin(\omega t)$,	Frequency = 50 Hz A = 1
Sag	$f(t) = A(1-a(u(t-t_1)-u(t-t_2)))\sin(\omega t)$,	$0.1\leqslant a\leqslant 0.9$; $T\leqslant t_2 - t_1\leqslant 9T$
Swell	$f(t) = A(1 + a(u(t-t_1)\text{-}u(t\text{-}t_2)))\sin(\omega t)$,	$0.1\leqslant a\leqslant 0.8$; $T\leqslant t_2 - t_1\leqslant 9T$
Interruption	$f(t) = A(1-a(t\sin(\omega t)$,	$0.1\leqslant a\leqslant 0.9$; $T\leqslant t_2 - t_1\leqslant 9T$
Flicker	$f(t) = A(1 + \alpha_f\sin(\beta_f \omega t))\sin(\omega t)$,	$\alpha_f = 0.1-0.2$ $\beta_f = 5-10$ Hz
Harmonic	$f(t) = A(\alpha_1.\ \sin(\omega t) + \alpha_3\sin(3\omega t) + \alpha_5.\ \sin(5\omega t) + \alpha_7.\ \sin(7\omega t))$	$0.05\leqslant \alpha_3 \leqslant 0.15$ $0.05\leqslant \alpha_5 \leqslant 0.15$ $0.05\leqslant \alpha_7 \leqslant 0.15$ $\sum \alpha_i^2 = 1$
Transient	$f(t) = A\sin(\omega t) + \alpha_{OSC}.\ \exp\left(-\left(\frac{t-t_1}{T_{OSC}}\right)\right).\ \sin\omega_{nosc}(t-t_1))$,	$T_{osc} = 0.008-0.04s$ $\omega_{nosc} = 100-400$ Hz
Swell with Harmonic	$f(t) = A(1 + \alpha(u(t-t_1)-u(t-t_2))).\ \alpha_1\sin(\omega t) + \alpha_3.\ \sin(3\omega t) + \alpha_5.\ \sin(5\omega t) + \alpha_7.\ \sin(7\omega t))$,	$0.1\leqslant \alpha \leqslant 0.8$ $T\leqslant t_2-t_1 \leqslant 9T$ $0.05\leqslant \alpha_i \leqslant 0.15$ $\sum \alpha_i^2 = 1$
Sag with harmonics	$f(t) = A(1-a(u(t-t_1)-u(t-t_2)))*A(\alpha_1.\ \sin(\omega t) + \alpha_3\sin(3\omega t) + \alpha_5.\ \sin(5\omega t) + \alpha_7.\ \sin(7\omega t))$,	$0.1\leqslant a\leqslant 0.9$; $T\leqslant t_2 - t_1\leqslant 9T$ $0.05\leqslant \alpha_3 \leqslant 0.15$ $0.05\leqslant \alpha_5 \leqslant 0.15$ $0.05\leqslant \alpha_7 \leqslant 0.15$ $\sum \alpha_i^2 = 1$
Flicker with harmonics	$f(t) = A(1 + \alpha_f\sin(\beta_f \omega t))\sin(\omega t)*A(\alpha_1.\ \sin(\omega t) + \alpha_3\sin(3\omega t) + \alpha_5.\ \sin(5\omega t) + \alpha_7.\ \sin(7\omega t))$,	$\alpha_f = 0.1-0.2$ $\beta_f = 5-10$ Hz $0.05\leqslant \alpha_3 \leqslant 0.15$ $0.05\leqslant \alpha_5 \leqslant 0.15$ $0.05\leqslant \alpha_7 \leqslant 0.15$ $\sum \alpha_i^2 = 1$
Interruption with harmonics	$f(t) = A(1-a(t\sin(\omega t)*A(\alpha_1.\ \sin(\omega t) + \alpha_3\sin(3\omega t) + \alpha_5.\ \sin(5\omega t) + \alpha_7.\ \sin(7\omega t))$,	$0.1\leqslant a\leqslant 0.9$; $T\leqslant t_2 - t_1\leqslant 9T$ $0.05\leqslant \alpha_3 \leqslant 0.15$ $0.05\leqslant \alpha_5 \leqslant 0.15$ $0.05\leqslant \alpha_7 \leqslant 0.15$ $\sum \alpha_i^2 = 1$
Spikes	$f(t) = \sin(\omega t) + sign(\sin(\omega t))* \sum k*(u(t-(t_1 + 0.02n))-(u(t-(t_2-0.02n)))$,	$0.1\leqslant k \leqslant 0.4, 0 \leqslant t_1$ $t_2\leqslant 0.5T$, $0.01T \leqslant t_2 - t_1\leqslant 0.05$ T

Table 2
Oscillatory Transients Classification.

Oscillatory transient	Spectral content	Typical duration	Typical voltage magnitude
Low frequency	< 5 KHz	0.3–50 ms	0–4 pu
Medium frequency	5–500 KHz	20 μs	0–8 pu
High frequency	0.5–5 MHz	5 μs	0–4 pu

Table 3
Impulsive Transients Categories.

Impulsive transient	Spectral content	Typical duration
Nanosecond	5 ns rise	Less than 50 ns
Microsecond	1 us rise	50 ns to 1 ms
Millisecond	0.1 ms rise	More than 1 ms

in arrangement of PQ aggravations.

Sundaram and Neela [45] presented a method based on Kalman filter and ANN for classification of various disturbances. The disturbance waveforms were generated through Matlab Simulink on a test system. Disturbances are inclusive of notch and flicker. The PQ features such as standard deviation, peak value and variances were extracted through Kalman filter. Rudhra and Gopikrishnan [37] proposed a generalized regression neural network for classification of PQ disturbances. Lopez-Ramirez et al. [25] proposed a model for classifying PQD through parameter estimation. Abdoos et al. [2] exhibited an advanced method for PQ aggravations identification in electrical power frameworks. Montoya et al. [31] made a survey of researches regarding the matter of energy quality distributed amid the period from 1970 to 2013. Li et al. [23] showed another method for discovery and grouping of PQ unsettling influences in light of the mix of S-transform and support vector machine.

Thirumala et al. [46] suggested another robotized acknowledgment approach in view of tunable-Q wavelet transform (TQWT) and a double multiclass SVM (MSVM) for recognition of power quality aggravations. Ma et al. [27] suggested a profound learning-based strategy brought into characterization of power quality disturbances (PQDs). Mahela and

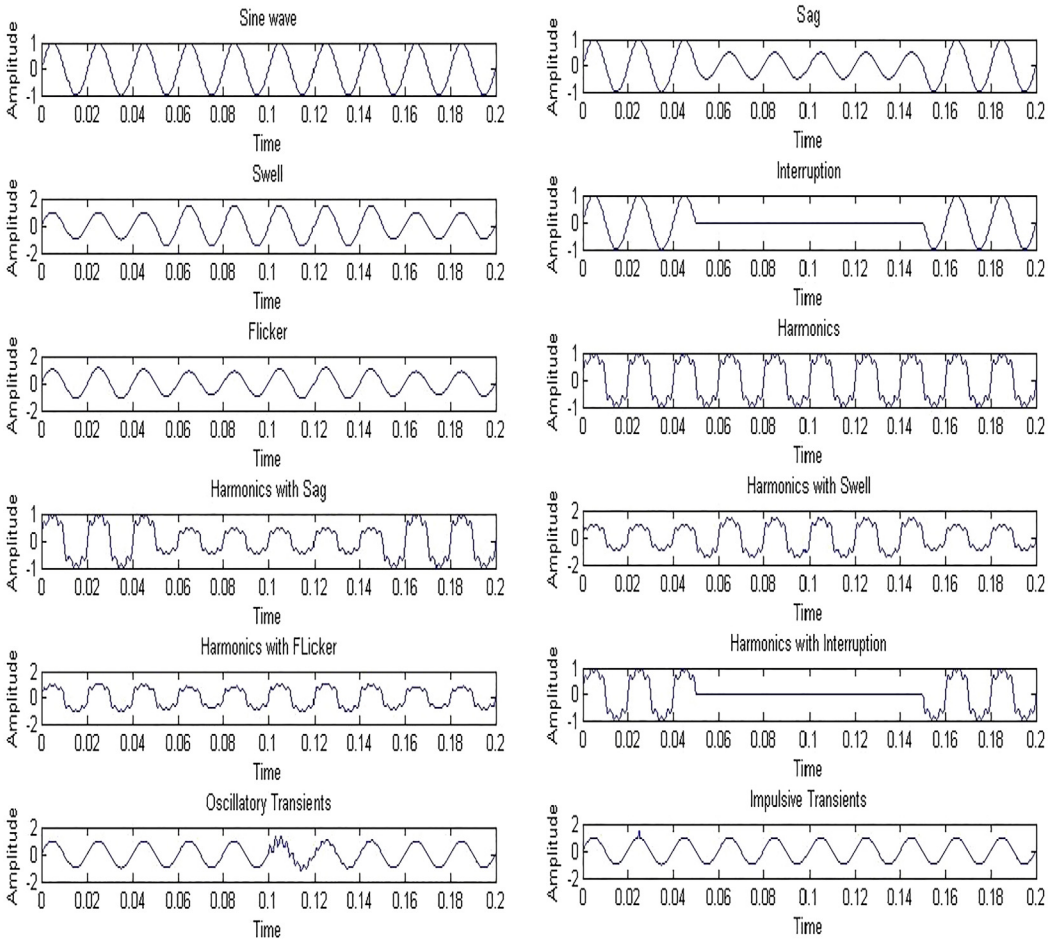


Fig. 3. All tested events.

Shaik [28] presented a method to perceive the power quality unsettling influences related with sun based vitality infiltration in distributed network. Morales-Velazquez et al. [32] presented a smart network based on field programmable gate array in residential home, industrial installation, and public building. The developed network computes features such as effective values, power factor, total harmonic distortion, etc. to detect power quality disturbances such as dips, swells, or interruptions.

Reddy and Sodhi [36] suggested a Stockwell transform. Das et al. [5] presented a novel procedure to analyze power quality unsettling influence cases in both single and additionally event. Kapoor et al. [18] recognized power quality based on Histogram of oriented gradients and SVM. Besides, there are other types of machine learning methods that could be applied to this suggested method in [42,41,43,44,3,11,13,26,33,7,48,47,22].

As shown above, significant works have been carried out for detecting time domain and frequency behaviors of events. The major issues in all these works are the accuracy of detection and complexity of the algorithms [20]. Also, we realize that a lot more needs to be done to rectify these issues by taking suitable action. Some algorithms like in [18] do not address the issues of cross over interferences which diminishes speeding up of the computational time of the method.

3. Proposed methodology

3.1. Framework

In this section, we propose a novel approach using KL Divergence and Standard deviation to detect events and SVM for event classification. The KL Divergence and Standard Deviation are calculated for each cycle of a given signal to know the exact order of harmonic present in the signal (Fig. 1).

The KL Divergence is used for calculation of probabilities of a particular cycle that is a member of two or more than two events. The most significant point of using KL Divergence is that it does not possess inequality. Thus, this research calculates KL Divergence for each cycle in comparison between given signal and Sine wave. Let us take an example of using KL Divergence in thermodynamics. KL Divergence measures thermodynamic availability. It is also useful even if we have only clues about experimental measurements. It describes distance to equilibrium or the amount of available work when multiplied by ambient temperature. Likewise, an example of standard deviation can be seen from the average height of men in a country. If the standard deviation is zero then all men would be exactly 70″. If it is 20″ then some men would be taller or shorter than the average. By this, standard deviation is used to calculate the quantity of difference of a set of data values.

The procedure of the proposed PQ identification comprises of two stages: detection and classification. In the first step, info flag is caught by observing gadgets. The essential highlights are then removed from the flag which is utilized for the classification of occasions. The classification step takes the removed highlights as information and the occasion classification is completed utilizing Support Vector Machine (SVM). Preprocessing venture of deposing the information flag is finished utilizing for discovery of occasion event. This task is performed to stay away from the antagonistic influence of clamor. To manage non-stationary signs (e.g. PQ occasion flag), numerous examination works are done. Endeavors are made to beat the impediments of Fourier

Table 4
Detection data for all the events for each cycle of events.

Events	Cycle No.	1	2	3	4	5	6	7	8	9	10
Sine	KL Divergence	0	0	0	0	0	0	0	0	0	0
	Standard deviation	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074
Sag	KL Divergence	0	0	0.9835	1.9628	1.9628	1.9628	1.9628	0.9801	0	0
	Standard deviation	0.7074	0.7074	0.4569	0.1768	0.1768	0.1768	0.1768	0.4569	0.7074	0.7074
Swell	KL Divergence	0	0	0.1220	0.2454	0.2454	0.2454	0.2454	0.1233	0	0
	Standard deviation	0.7074	0.7074	0.9795	1.2379	1.2379	1.2379	1.2379	0.9795	0.7074	0.7074
Interruption	KL Divergence	0	0	1.7644	3.9479	3.9479	3.9479	3.9479	1.4716	0	0
	Standard deviation	0.7074	0.7074	0.3857	0	0	0	0	0.3857	0.7074	0.07074
Flicker	KL Divergence	0.0196	0.0386	0.0093	0.5062	0.0901	0.0196	0.0386	0.0093	0.5061	0.0901
	Standard deviation	0.7862	0.8346	0.7078	0.5803	0.6291	0.7862	0.8346	0.7078	0.5803	0.6291
Harmonics (3rd, 5th & 7th)	KL Divergence	0.2373	0.2373	0.2373	0.2373	0.2373	0.2373	0.2373	0.2373	0.2373	0.2373
	Standard deviation	0.7309	0.7309	0.7309	0.7309	0.7309	0.7309	0.7309	0.7309	0.7309	0.7309
Harmonics (only 3rd)	KL Divergence	0.0859	0.0859	0.0859	0.0859	0.0859	0.0859	0.0859	0.0859	0.0859	0.0859
	Standard deviation	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214
Harmonics (only 5th)	KL Divergence	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680
	Standard deviation	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214
Harmonics (only 7th)	KL Divergence	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014	0.2014
	Standard deviation	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214	0.7214
Harmonics with sag	KL Divergence	0.2373	0.2373	1.1158	1.9880	1.9880	1.9880	1.9880	1.1095	0.2373	0.2373
	Standard deviation	0.7309	0.7309	0.4632	0.1827	0.1827	0.1827	0.1827	0.4632	0.7309	0.7309
Harmonics with swell	KL Divergence	0.2373	0.2373	0.5020	0.7678	0.7678	0.7678	0.7678	0.5032	0.2373	0.2373
	Standard deviation	0.7309	0.7309	1.0079	1.2790	1.2790	1.2790	1.2790	1.0079	0.7309	0.7309
Harmonics with flicker	KL Divergence	0.3364	0.3273	0.2956	1.0510	0.5896	0.3364	0.3273	0.2956	1.0510	0.5897
	Standard deviation	0.8120	0.8617	0.7313	0.6001	0.6503	0.8120	0.8617	0.7313	0.6001	0.6503
Harmonics with interruption	KL Divergence	0.2373	0.2373	1.7126	3.9479	3.9479	3.9479	3.9479	1.5889	0.2373	0.2373
	Standard deviation	0.7309	0.7309	0.3796	0	0	0	0	0.3796	0.7309	0.7309
Osc. transient	KL Divergence	0	0	0	0	0	0.0319	0.0066	0.0051	0	0
	Standard deviation	0.7074	0.7074	0.7074	0.7074	0.7074	0.7887	0.7189	0.7085	0.7074	0.7074
Impulsive transient	KL Divergence	0	0	0	0	0	0	0	0	0	0
	Standard deviation	0.7074	0.7240	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074	0.7074

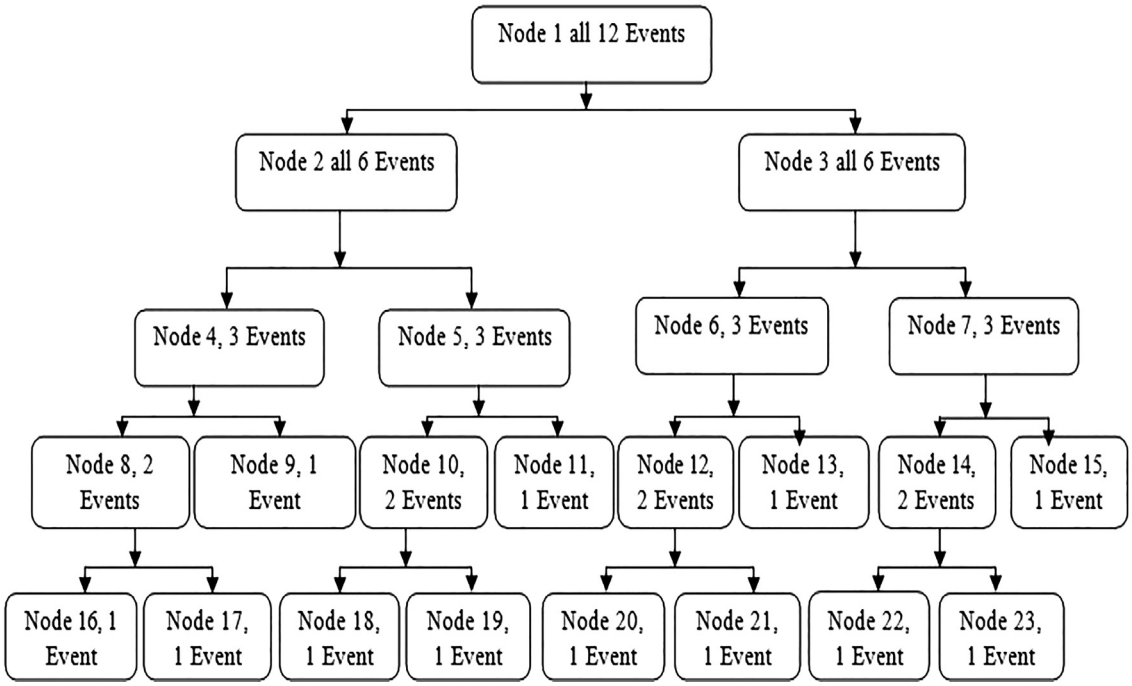


Fig. 4. Classification results.

Table 5
SVM training dataset of Sag [18].

%	KL Divergence	Standard deviation
10	0.0397	0.6366
20	0.8355	0.5657
30	0.7742	0.4952
0	1.3715	0.4244
50	1.2949	0.3537

Table 6
SVM training dataset of Swell [18].

%	KL Divergence	Standard deviation
10	0.0155	0.7781
20	0.0466	0.8489
30	0.0809	0.9196
40	0.1225	0.9903
50	0.1577	1.0611

Table 7
SVM training dataset of Interruption [18].

%	KL Divergence	Standard deviation
90	2.9479	0.0707
92	2.9479	0.0566
94	2.9479	0.0424
96	2.9479	0.0283
100	3.9479	0

Table 8
SVM training dataset of Flicker [18].

KL Divergence	Standard deviation
0.0122	0.7665
0.0244	0.8028
0.0051	0.7076
0.0959	0.6121

Table 9
SVM training dataset of Oscillatory Transient [18].

Freq.	KL Divergence	Standard deviation
100	0.0325	0.7486
200	0.1059	0.7448
300	0.0449	0.7126
400	0.0245	0.7446

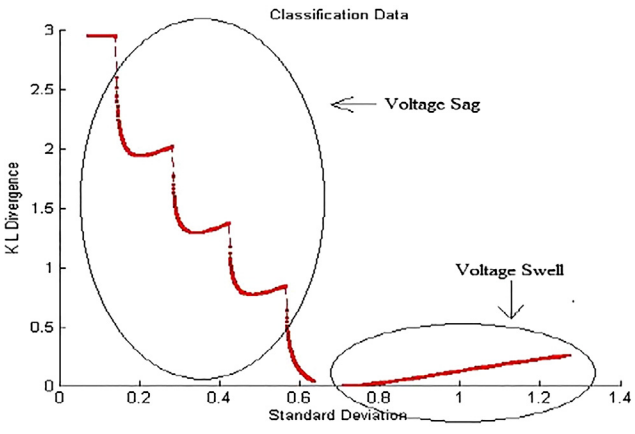


Fig. 5. Classification data for Voltage Sag and Voltage Swell.

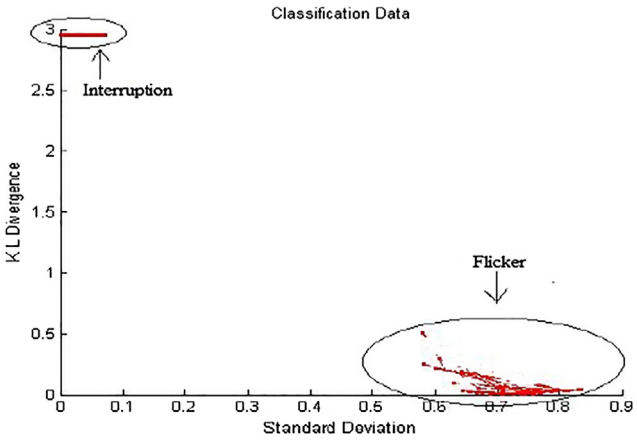


Fig. 6. Classification data for Interruption and Flicker.

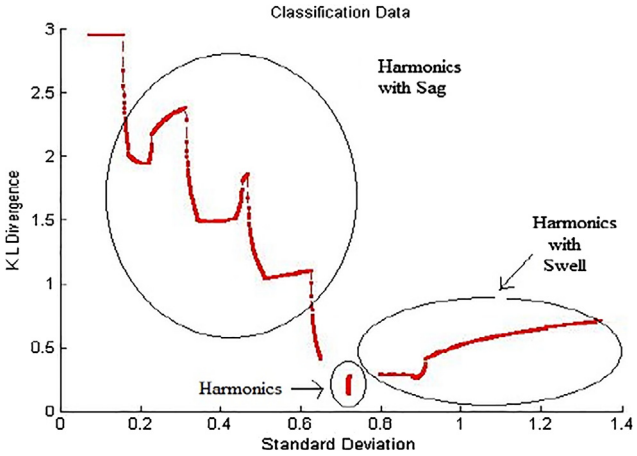


Fig. 7. Classification data for Harmonics, Harmonics with Sag and Harmonics with Swell.

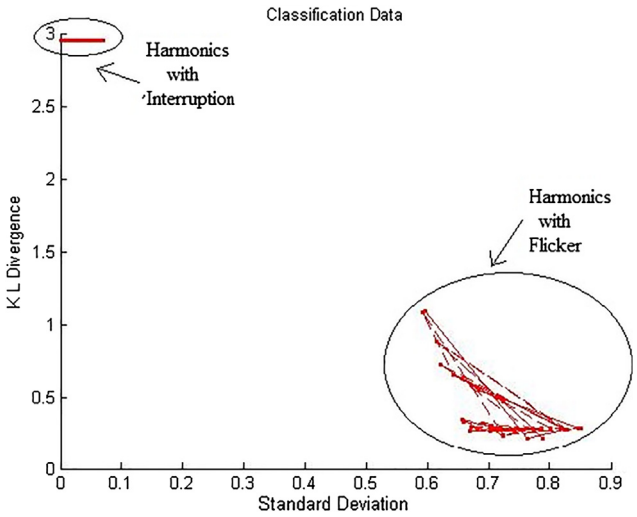


Fig. 8. Classification data for Harmonics with interupption and Harmonics with Flicker.

change that expect the flag to be stationary. Brief time Fourier transforms (STFT) and wavelet transforms (WT) are few cases of such changes. Nonetheless, determination of windows is a noteworthy issue in STFT which is wiped out in WT [8]. WT tries to distinguish both the scale and space data of the occasion which make it a compelling apparatus for evaluation and investigation of PQs. Utilizing WT and multi

Table 10
Comparisons of ST, EMD-HT using different events.

Event	ST	EMD-HT	Proposed method
Sag	NO	NO	YES
Swell	NO	NO	YES
Flicker	YES	YES	YES
Interruption	YES	YES	YES
Harmonics	YES	YES	YES
Harmonics with sag	YES	YES	YES
Harmonics with swell	YES	YES	YES
Oscillatory transients	YES	YES	YES
Impulsive transients	YES	YES	YES
Notches	YES	YES	NO
Order of harmonics	NO	NO	YES

Table 11
Comparison of existing and proposed work in term of accuracy.

No.	Authors	Method	Accuracy (%)
1.	Abdoos et al. [2]	WT + SVM	92.33
2.	Ma et al. [27]	WT + RVM	93.47
3.	Li et al. [23]	WT + NN	89.92
4.	Kapoor et al. [18]	HOG + SVM	93.87
5.	Proposed Work	KL Divergence + Standard Deviation	94.02

determination investigation strategy flag is deteriorated up to a fixed level of determination. An extensive variety of mother wavelets are accessible to look over and each of the have novel properties. For PQ applications, it has been accounted to choose wavelet ideally [24].

3.2. Standard Deviation

Standard Deviation (SD) is a metric system to calculate the quantity of difference or distribution of a set of data values. It is a square root of variance for measuring uncertainty. Consider R being a random variable with mean values μ. We have the formula in Eq. (1).

$$E_v[R] = \mu. \tag{1}$$

E_v is the average/expected value of R. The SD of R is the quantity in Eq. (2).

$$\begin{aligned}\sigma &= \sqrt{Ev[(R-\mu)^2]},\\ &= \sqrt{Ev[R]^2 + Ev[(-2\mu R)] + Ev[\mu^2]},\\ &= \sqrt{Ev[R]^2 - 2\mu Ev[R] + \mu^2},\\ &= \sqrt{Ev[R]^2 - (Ev[R])^2}.\end{aligned} \tag{2}$$

3.3. KL Divergence

Kullback–Leibler (KL) divergence is defined in Eq. (3) as logarithmic difference between P_n and Q_n. Whenever $P_n(i) = 0$, the i-th term is zero as shown in Eq. (4).

$$D_{KL}(P||Q) = \sum_i P_n(i)\ln\frac{P_n(i)}{Q_n(i)}, \tag{3}$$

$$\lim_{x\to 0} x\ln(x) = 0. \tag{4}$$

In continuous domain, the Kullback-Leibler divergence is defined as:

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} P_n(i)\ln\frac{P_n(i)}{Q_n(i)}, \tag{5}$$

where p_n and q_n denote the densities of P_n and Q_n. The KL divergence is non-negative.

$$D_{KL}(P||Q) \geqslant 0. \tag{6}$$

3.4. Classifier

Classifier is used to classify events on the basis of characteristics to which classes they belong. In this research, we use Support Vector Machine (SVM) to determine a decision boundary between pattern classes which minimizes the error level.

It finds out best separation between patterns and provides higher accuracy in classification [35]. There is another type of SVM which is non-linear polynomial and radial basis function to map the sample to get the higher dimensional space. The basic methodology of SVM classifier is shown in Fig. 2.

3.5. Advantages of the proposed method

The proposed method has two steps: detection and classification involving four parts: preprocessing, KL Divergence, Standard Deviation and Classification. At first, preprocessing is carried out for incoming events to reduce noises. Once the incoming events are successfully preprocessed, we compute KL Divergence and Standard Deviation for each cycle of aforementioned events. After finding the KL Divergence and Standard Deviation, we compare the obtained values with the values of sine wave; if they differ then they are applied to the SVM classifier for classification purpose. The advantages of this method over previous ones are: high accuracy, not loose distinguishing characteristics, etc.

4. Experiments

4.1. Environmental set up

Tables 1 and 2 describe a list of power quality events and oscillatory transient.

Impulsive transients can be subdivided into three categories as given in Table 3.

The following sections present simulation results of detection and classification of signal disturbances.

4.2. Detection

KL Divergence and Standard Deviation are calculated for each cycle of a given signal and then compared with those of Sine wave. If there is any difference in the values of both the parameters then disturbance in the desired signal is present. In the given Fig. 3, we can easily observe that whenever disturbance exists then there is a variation in the amplitude of the signal used for calculation of both parameters. From Table4, it is clear that both the parameters deviates from the original value whenever disturbance. If SVM is trained with data of harmonics with order of 3, 5, 7, etc. then this method can be used to know the exact order of harmonic present in the signal.

4.3. Classification

Fig. 4 shows the results of classification using SVM.

Table 5 depicts the results of KL divergence and Standard Deviation by training of SVM classifier with Sag as a sample. Table 6 depicts the results of KL divergence and Standard Deviation by training of SVM classifier with Swell as a sample. Likewise, Tables 7–9 are with Interruption, Flicker and Oscillatory Transient.

Values for Ideal Sine Cycle are 0 and 0.7074 for KL Divergence and Standard Deviation respectively. In classification stage, classification of 9 events is shown in Figs. 5–8. A similar process for classification has been done with the empirical-mode decomposition and Hilbert transform-based method (EMD-HT) and ST method [4]. The comparison has been shown in Table 10. In Figs. 5–8, different samples of various events have been taken and then value sof Standard Deviation and KL Divergence has been obtained. For voltage sag, 800 values have been

taken. Likewise, voltage swell consists of 900 values; voltage flicker has 500 values; harmonics has 1000 values; and harmonics with sag and harmonics with Swell have 800 values. It is clear that after calculating values of Standard Deviation and KL divergence, events can be detected and classified. For some events, it is difficult to classify because those values for two events are very close to each other or they may overlap. Like in case of Flicker and Voltage Swell, it is possible that classification may not be very effective and also in case of interruption and harmonics with interruption, false classification may take place.

4.4. Main findings

1. When a straight load is available to the framework, it draws a sinusoidal current where the recurrence and voltages are same. In any case, with apathetic stage, the subsequent current waveform might be exceptionally mind boggling in nature. Consequently, sounds are produced in light of the nearness of non-direct electric loads that are associated with the framework. Most basic cases of gadgets that go about as non-direct loads incorporate basic office gear.
2. Voltage swell is expansion in the ostensible rms voltage between 1.1 p.u. and 1.8 p.u.
3. Voltage droop is the abatement in the ostensible rms voltage between 0.1 p.u. and 0.9 p.u.
4. Time of voltage swell is from 0.5 cycles to 1 min.
5. Voltage intrusion happens because of free association, real blames and reclosing of electrical switch.
6. Size of oscillatory transient is between 1.2 and 1.5 p.u.
7. Placing quick acting tap changers can moderate voltage swell.

In conclusion, voltage swell can be partitioned into three expansive classes in view of the term of interim for which swell exists. These classes are immediate, flashing and transitory swell if there should be an occurrence of voltage hang. Table 11 indicates the comparison results in term of accuracy. We can see that Abdoos et al. [2] used Pure signal, sag, swell, interruption, harmonics, transient, sag and harmonics, swell and harmonics, flicker to yield out the accuracy of 92.33%. Abdoos et al. [2] havenot done classification due to which fault events have also been included. However, Ma et al. [27] and Li et al. [23] used the same but they also did classification. However their approach yielded out an accuracy of 93.47% and 89.92% respectively. In Kapoor et al. [18], HOD and SVM have been used, but the processing time is high. However, they were successful to achieve accuracy of 93.87%.

5. Conclusion

In this paper, we proposed a new method to identify power quality events by KL Divergence and Standard Deviation. It is clearly visible that KL Divergence along with Standard Deviation can be effectively used for detection. It is more effective than other techniques used in the past because this technique can be applied to single cycle of the input signal whereas the previous ones require multiple cycles. Therefore, the previous works are complex and time inefficient and require more memory for storing more number of cycles. These problems are reduced in by using KL Divergence and Standard Deviation. Those techniques can be used to find the order of harmonic present in the signal if SVM is trained with suitable data which have not been done in the past references.

The new method is absolutely significant in both theory and practice. Practically, both researchers and industries focused on consuming lesser amount of memory which can be achieved using KL Divergence and Standard Deviation. Classification of the events using SVM has yielded out the utilization of high memory in a very significant way. In the other words, utilization has been very low in comparison to the previous works. Ordering of harmonics plays an important role in power quality detection. Though researchers have used this in swelling and sagging, using KL Divergence technique along with Standard Deviation, we have found the order of harmonic present in the signal. This not only detects the power quality in a significant way, but also reduces the processing time and decreasing the complexity with properly maintaining the order of harmonics.

Future works of this research will include the following things: a deep learning method for classification of disturbances, optimal feature selection for PQDs, particle swarm optimization for classifying PQDs, and cross-validation with respect to various schemes.

References

[1] A. Abdelazeem, A. Abdelsalam, A. Eldesouky, A. Sallam, Characterization of power quality disturbances using hybrid technique of linear kalman filter and fuzzy expert system, Electr. Power Syst. Res. 83 (2012) 41–50.
[2] A.A. Abdoos, P.K. Mianaei, M.R. Ghadikolaei, Combined VMD-SVM based feature selection method for classification of power quality events, Appl. Soft Comput. 38 (2016) 637–646.
[3] M. Ali, L.H. Son, N.D. Thanh, N. Van Minh, A neutrosophic recommender system for medical diagnosis based on algebraic neutrosophic measures, Appl. Soft Comput. (2018), http://dx.doi.org/10.1016/j.asoc.2017.10.012.
[4] B. Biswal, S. Mishra, R. Jalaja, Automatic classification of power quality events using balanced neural tree, IEEE Trans. Ind. Electron. 61 (1) (2014) 521–530.
[5] S. Das, A.K. Pradhan, A. Kedia, S. Dalai, B. Chatterjee, S. Chakravorti, Diagnosis of power quality events based on detrended fluctuation analysis, IEEE Trans. Ind. Electron. (2018) 24–31.
[6] P.K. Dash, B.K. Panigrahi, G. Panda, Power quality analysis using S-transform, IEEE Trans. on Power Delivery 18 (2) (2003) 406–411.
[7] C.N. Giap, L.H. Son, F. Chiclana, Dynamic structural neural network, J. Intell. Fuzzy Syst. (2018) 1–12.
[8] M. Gouda, M.M.A. Salama, M.R. Sultan, A.Y. Chikhani, Power quality detection and classification using wavelet multiresolution signal decomposition, IEEE Trans. Power Delivery 14 (4) (1999) 1469–1476.
[9] Y. Gu, M. Bollen, Time frequency and time scale domain analysis of voltage disturbances, IEEE Trans. Power Delivery 15 (4) (2000) 1279–1284.
[10] M. Gupta, R. Kumar, Detection of power quality disturbances using symbolic dynamics, Recent Advances and Innovations in Engineering (ICRAIE), 2014, IEEE, 2014, pp. 1–8.
[11] D.T. Hai, L.H. Son, T. Le Vinh, Novel fuzzy clustering scheme for 3D wireless sensor networks, Appl. Soft Comput. 54 (2017) 141–149.
[12] M. Hajian, A.A. Foroud, A new hybrid pattern recognition scheme for automatic discrimination of power quality disturbances, Measurement 51 (2014) 265–280.
[13] D.J. Hemanth, J. Anitha, L.H. Son, Brain signal based human emotion analysis by circular back propagation and deep kohonen neural networks, Comput. Electr. Eng. 68 (2018) 170–180.
[14] G.T. Heydt, et al., Applications of the windowed FFT to electric power quality assessment, IEEE Trans. Power Deliv. 14 (4) (1999) 1411–1416.
[15] S. Jamali, A.R. Farsa, N. Ghaffarzadeh, Identification of optimal features for fast and accurate classification of power quality disturbances, Measurement 116 (2018) 565–574.
[16] P. Janik, T. Lobos, Automated classification of power quality disturbances using SVM and RBF networks, IEEE Trans. Power Delivery 21 (3) (2006) 1663–1669.
[17] Jiaqi Li, M.V. Chilukuri, Power supply quality analysis using S-transform and SVM classifier, J. Power Energy Eng. 2 (2014) 438–447.
[18] R. Kapoor, R. Gupta, L.H. Son, S. Jha, R. Kumar, Detection of power quality event using histogram of oriented gradients and support vector machine, Measurement 120 (2018) 52–75.
[19] S. Khokhar, A.A.B.M. Zin, A.S.B. Mokhtar, M. Pesaran, A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances, Renew. Sustain. Energy Rev. 51 (2015) 1650–1663.
[20] S. Khokhar, A.A.M. Zin, A.P. Memon, A.S. Mokhtar, A new optimal feature selection algorithm for classification of power quality disturbances using discrete wavelet transform and probabilistic neural network, Measurement 95 (2017) 246–259.
[21] R. Kumar, B. Singh, D.T. Shahani, Recognition of single-stage and multiple power quality events using Hilbert-Huang transform and probabilistic neural network, Electr. Power Compon. Syst. 43 (6) (2015) 607–619.
[22] T. Le, B. Vo, S.W. Baik, Efficient algorithms for mining top-rank-k erasable patterns using pruning strategies and the subsume concept, Eng. Appl. Artif. Intell. 68 (2018) 1–9.
[23] J. Li, Z. Teng, Q. Tang, J. Song, Detection and classification of power quality disturbances using double resolution S-transform and DAG-SVMs, IEEE Trans. Instrum. Meas. 65 (10) (2016) 2302–2312.
[24] C.H. Lin, M.C. Tsao, Power quality detection with classification enhancible wavelet-probabilistic network in a power system, IEEE Proc. Gener. Transm. Distrib. 152 (2005) 969–976.
[25] Lopez-Ramirez, et al., EMD-based feature extraction for power quality disturbance classification using moments, Energies 09 (07) (2016) 565–570.
[26] A. Louati, L.H. Son, H. Chabchoub, Smart routing for municipal solid waste collection: a heuristic approach, J. Ambient Intell. Hum. Comput. (2018) 1–20.
[27] J. Ma, J. Zhang, L. Xiao, K. Chen, J. Wu, Classification of power quality disturbances via deep learning, IETE Tech. Rev. 34 (4) (2017) 408–415.
[28] O.P. Mahela, A.G. Shaik, Power quality recognition in distribution system with solar

energy penetration using S-transform and Fuzzy C-means clustering, Renewable Energy 106 (2017) 37–51.
[29] S. Mishra, C.N. Bhende, B.K. Panigrahi, Detection and classification of power quality disturbances using S-transform and probabilistic neural network, IEEE Trans. Power Delivery 23 (2008) 280–287.
[30] S.R. Mohanty, N. Kishor, P.K. Ray, J.P. Catalo, Comparative study of advanced signal processing techniques for islanding detection in a hybrid distributed generation system, IEEE Trans. Sustain. Energy 6 (1) (2015) 122–131.
[31] F.G. Montoya, A. García-Cruz, M.G. Montoya, F. Manzano-Agugliaro, Power quality techniques research worldwide: a review, Renew. Sustain. Energy Rev. 54 (2016) 846–856.
[32] L. Morales-Velazquez, R. de Jesus Romero-Troncoso, G. Herrera-Ruiz, D. Morinigo-Sotelo, R.A. Osornio-Rios, Smart sensor network for power quality monitoring in electrical installations, Measurement 103 (2017) 133–142.
[33] R.T. Ngan, M. Ali, L.H. Son, δ-equality of intuitionistic fuzzy sets: a new proximity measure and applications in medical diagnosis, Appl. Intelligence 48 (2) (2018) 499–525.
[34] R. Polikar, The Engineer's Ultimate Guide to Wavelet analysis, The Wavelet Tutorial, 1999, pp. 25–46.
[35] F. Qi, C. Bao, Y. Liu, A novel two-step SVM classifier for voiced/unvoiced/silence classification of speech, J. Power Energy Eng. 2 (4) (2004) 438–447.
[36] M.V. Reddy, R. Sodhi, A modified S-transform and random forests-based power quality assessment framework, IEEE Trans. Instrum. Meas. 67 (1) (2018) 78–89.
[37] S. Rudhra, M. Gopikrishnan, Identification of power quality issues using Kalman filter, Ind. J. Sci. Eng. Appl. Sci. 08 (32) (2016) 1–5.
[38] L.C. Saikia, S.M. Borah, S. Pait, Detection and classification of power quality disturbances using wavelet transform and neural network, in: IEEE Annual India Conference, 2010, pp. 78–84.
[39] C. Sankaran, Power Quality, CRC Press, 2017.
[40] S. Santoso, W.M. Grady, E.J. Powers, J. Lamoree, S.C. Bhatt, Characterization of distribution power quality events with fourier and wavelet transforms, IEEE Trans. Power Delivery 15 (2000) 247–254.
[41] L.H. Son, Measuring analogousness in picture fuzzy sets: from picture distance measures to picture association measures, Fuzzy Optim. Decis. Making 16 (2017) 359–378.
[42] L.H. Son, P.H. Phong, On the performance evaluation of intuitionistic vector similarity measures for medical diagnosis, J. Intell. Fuzzy Syst. 31 (3) (2016) 1597–1608.
[43] L.H. Son, T.M. Tuan, Dental segmentation from X-ray images using semi-supervised fuzzy clustering with spatial constraints, Eng. Appl. Artif. Intell. 59 (2017) 186–195.
[44] L.H. Son, P.V. Viet, P.V. Hai, Picture inference system: a new fuzzy inference system on picture fuzzy set, Appl. Intelligence 46 (3) (2017) 652–669.
[45] Sundaram, Neela, Analysis of power quality disturbances based on Kalman filter and MLP neural network, ARPN J. Eng. Appl. Sci. 11 (9) (2016) 5676–5687.
[46] K. Thirumala, S.P. Maganuru, T. Jain, A. Umarikar, Tunable-Q wavelet transform and dual multiclass SVM for online automatic detection of power quality disturbances, IEEE Trans. Smart Grid (2016), http://dx.doi.org/10.1109/TSG.2016.2624313.
[47] B. Vo, T. Le, G. Nguyen, T.P. Hong, Efficient algorithms for mining erasable closed patterns from product datasets, IEEE Access 5 (2017) 3111–3120.
[48] B. Vo, T. Le, W. Pedrycz, G. Nguyen, S.W. Baik, Mining erasable itemsets with subset and superset itemset constraints, Expert Syst. Appl. 69 (2017) 50–61.
[49] P.S. Wright, Short time fourier transform and wigner-ville distributions applied to the calibration of power frequency harmonic analyzers, IEEE Trans. Instrum. Meas. (1999) 475–478.
[50] D. Yong, S. Bhowmik, F. Magnago, An effective power quality classifier using wavelet transform and support vector machines, Expert Syst. Appl. 42 (15–16) (2015) 6075–6081.
[51] F.Z. Zhao, R.G. Yang, Power-Quality Disturbance Recognition Using S-Transform, IEEE Trans. Power Del. 22 (2) (2007) 944–950.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Cement Industry and Strategies for Mitigating Carbon Emissions

Ahana Ghosh[1], Anubha Mandal[2]
[1]Environmental Engineering, Delhi Technological University, New Delhi, India -110039
[2]Environmental Engineering, Delhi Technological University, New Delhi, India- 110039

Abstract**—** *Cement is the largest mass manufactured man made product on earth. The demand for cement is on a continual rise, as more and more developing countries strive for better infrastructure. This demand has, however, entailed an unacceptable increase in the carbon emissions as the cement manufacturing industry is one of the most carbon releasing industries in the world; responsible for a large part of the global carbon emissions. The dangerously high levels of Carbon Dioxide have contributed to a large scale climate change which has global repercussions. The need of the hour is an effective yet inexpensive mechanism to trim down the carbon emissions from the cement factories. In this paper, the main industrial as well as the governmental strategies for alleviating the carbon emissions of the cement industry are reviewed, focusing on the carbon taxation for the latter. This review has observed a comprehensive literature in term of the peer reviewed journals, research papers, industry reports, authentic websites etc on the cement industry and the strategies to reduce the carbon emissions.*

Keywords**—** *Cement: carbon: carbon tax: environment: industries: strategies.*

I. INTRODUCTION

The manufacture of most industrial materials has some form of impact on the environment. Research is being conducted to reduce this impact and promote sustainable development.
One such industrial material which has a substantial effect on the environment, specifically on the carbon levels in the atmosphere is the vastly manufactured and consumed raw material: cement.

Currently, the cement industry accounts for almost 5%-7% [1] of the global CO_2 production. Over the recent years, cement production has witnessed an exponential increase in developing countries to meet the needs of a rapidly urbanizing civilization [2] and the carbon levels in the environment have seen a proportional increase. If nothing is done to control them, this ubiquitous industry will account for nearly 33% of the global carbon levels by 2050. [3]

This is indeed a cause for worry, as CO_2 is a greenhouse gas which when present in large quantities in the atmosphere can contribute to dangerous phenomena like global warming and climate change.[4] In such a situation, it is the need of the hour to find economically viable methods to help propagate a low CO_2 emitting cement industry.

It has been estimated that fossil fuel combustions account for a mere 30% of the amount of CO_2 generated during the production of cement whereas the calcination of the limestone accounts for almost 60% of the total carbon emitted.[5] This is both good as well as bad news. The carbon emissions cannot be controlled in the cement industry by simply increasing the energy efficiency; the problem must be tackled at the very base itself, by adjusting or changing the constituents of the cement, while keeping in mind the economic feasibility of the cement.

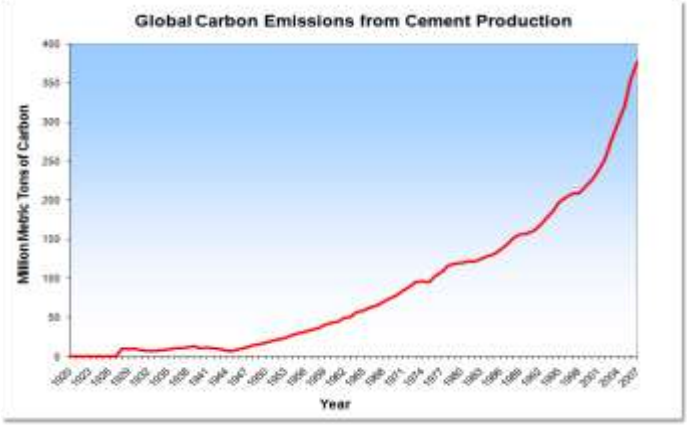


Fig. 1. Global carbon emissions from cement production (Boden, T. A., G. Marland, and R. J. Andres. 2010).

This paper gives an overview of the various major strategies that the governments as well as the cement industry have considered to lessen the amount of carbon released from the cement production process. The rest of the paper is organized as follows: The next section reviews the industrial plans and the subsequent section focuses on the governmental initiatives. After that, the conclusions are discussed and suggestions are made for the areas that merit further research.

II. INDUSTRIAL PLANS

If the carbon emissions are to be reduced, it is essential that the industries take initiative without any external pressure. For this, awareness as well as incentives by the government is necessary. Discussed below are some of the strategies applied by different cement companies to mitigate their emissions.

Carbon Capture and Storage

Carbon Capture and Storage (CCS) is one of the foremost technologies that are available in the industry today. It is the process of capturing the carbon from the source site (like a cement plant) and then transporting it to a storage facility and depositing it in a location where it will not impair the environment. Over the last two decades, the feasibility of this method has been researched upon considerably. [6] Research has shown that the expenditure of this strategy would cause in increase in the cost of cement production 2-3 times. [7] As

Ahana Ghosh and Anubha Mandal, "Cement industry and strategies for mitigating carbon emissions," *International Research Journal of Advanced Engineering and Science*, Volume 3, Issue 2, pp. 175-178, 2018.

such, it is yet to be proven for large scale use and is mainly suitable for those industries which do not have any other option to reduce their carbon footprint.

Use of Supplementary Cementitious Materials

These materials can be used to replace a certain percentage of the clinker that is used to manufacture the cement and hence reduce the carbon content of the clinker. These materials are mostly the by- products of other industries like fly ash, calcined clay, natural pozzolans or geopolymers and silica fume. Due to their advantages like cost effectiveness [2], long term durability [8] and ease of use, they have been used since the 1990s in the cement industry [9]. However, they come with their own set of disadvantages. The availability of these materials varies regionally [7] and the potential for their usage for most of these materials has already been explored. Hence there is little scope for further carbon reduction using these methods. Moreover, the biggest question that most research done till now fails to answer is that exactly upto how much percentage can the substitution be done without compromising on the durability of the cement? [10, 2] Furthermore, the cost analysis in terms of the labor required is still vague in most literature. For example, geopolymers are hailed to have widespread advantages like reduction of the CO2 emissions by 44-64% [11] over Ordinary Portland Cement, increased durability and better workability [12]. However, the disadvantages of these geopolymers like the fact that the making of the geopolymer concrete requires handling of the hazardous wastes and hence requires specialized training [13] and other technical difficulties were not accounted for in the cost analysis. Lastly, field studies have not yet proved conclusive for most of these materials and further research is needed to improve their usage in cement.

Use of Cement made from Alternative Clinkers

The main source of the carbon emission during cement manufacture is the stage at which the calcination of the CaO occurs. [14]. Hence, cement made from the alternative clinkers may allow for a substantial reduction in the carbon production. The paper by Gartner and Sui [15] provides an exhaustive analysis of the alternatives to Portland cement. Some of their conclusions are discussed below.

TABLE I. Comparison of alternative clinkers.

Cement	Advantage	Disadvantage
1] Belite rich Portland Cement	Carbon reduction upto 10% per unit clinker	Takes more time to gain strength as compared to OPC
2]Belite Calcium Sulfo-aluminate (CSA) Cements	Carbon reduction upto 20% per unit clinker.	The cost of the raw material is very high.
3]Magnesium Based Cements	Made from ultramafic rocks which have the inherent ability to capture carbon. Globally abundant raw materials.	As of now, no energy efficient industrial manufacturing process has been invented.

Use of Alternative Fuels

Fuels account for only 25-30% [16] of the carbon released during the cement manufacturing process and hence changing their compositions does not have a massive impact on the carbon emissions. The area does not have much more scope for exploitation [7] as a variety of alternative fuels have already been tried and tested and are currently being used in the cement industry.

III. GOVERNMENTAL POLICIES

Due to the difficulty faced by the cement industry in reducing the carbon emissions, the attention of the worldwide leaders and organizations as well as the governments of individual countries have riveted towards policies and strategies that could reduce the domestic greenhouse gases emissions. They can be very broadly divided into 3 types of approaches: - Voluntary approaches (VA), Trading and Carbon Taxes (CT). In this section, after a brief overview of VA and Trading, we shall focus on the discussion of Carbon Taxes.

Voluntary Approaches

VAs can be classified into one of four types: unilateral commitments by industry; private agreements between industry and stakeholders; environmental agreements negotiated between industry and government; voluntary programmes developed by government that individual firms can join. [17]. While this policy has been applauded for its flexibility and it's relatively less effect on the competitiveness of the companies which produce cement, it has also been criticized for the laxity in the coverage of the industries and the ineffectiveness of the implementation.

Carbon Trading

Carbon trading is a flexible mechanism introduced by the Kyoto Protocol, which limits the carbon emission from the industries by granting companies the permit to emit only a certain amount of carbon. This cap and trade mechanism, although effective to a certain degree, as proved by the studies of Shammin and Bullard (2009), is also a very complicated and expensive mechanism to implement. This system requires a completely new administrative system to aid the establishment of a competent trading market. [18] Moreover, the carbon trading mechanism merely shifts the production of carbon from one country to the other; which cannot be a permanent solution to the carbon problem as this issue is a global issue.

Carbon Taxes

Due to the vast carbon footprint of the cement industry, the governments of various countries have tried implementing a tax known as Carbon Tax in order to hold the cement companies answerable for the carbon that they generate. If set high enough, it becomes a potent financial incentive that motivates switches to clean energy across the economy, simply by making it more economically rewarding to move to less carbon intensive manufacturing methods. Carl and Fedor (2016) have come to the conclusions that an effective carbon tax with a good rate has been preferred over other governmental policies, especially since the last decade. Research has been done exhaustively in quite a few areas relating to the carbon tax. Certain conclusions from them are discussed below.

Ahana Ghosh and Anubha Mandal, "Cement industry and strategies for mitigating carbon emissions," *International Research Journal of Advanced Engineering and Science*, Volume 3, Issue 2, pp. 175-178, 2018.

Research done states that the different impacts of the carbon tax could arise from the fact that different rates are followed in different countries [19] and that the design of a proper carbon tax may be able to alleviate the negative impacts of its implementation. [20]. Moreover, the carbon tax can be considered an efficient system only when the tax is set high enough that companies have enough incentive to switch to a lower carbon intensive manufacturing process.[20]. However in China, an analysis done to evaluate the preference of companies to carbon tax found that companies prefer a low rate of carbon tax in the beginning; about 1 to 4 US dollars per tonne of CO_2 emitted.[21]. This, they claim, will reduce the carbon emissions and at the same time, not affect the competitiveness of the companies. A research done based on the Saudi Arabian cement industry suggests a compromise at 27 US dollars per ton of carbon [6] claiming that at this rate the profit of the industries would not be compromised and the emissions would significantly be reduced. Suggestions have also been made of starting at a low carbon tax rate and then gradually moving to high rates to avoid a sudden economic pressure on the industries. [22]

A survey done in the University of Geneva evaluates carbon taxes with respect to their competitiveness, distributional and environmental effects. [23]. Competitiveness indicates the ability of a company to sell its goods and services in the domestic as well as the global market. Distributional effects can be regressive i.e. the bulk of the tax falls more on the low income population or progressive i.e households with higher income pay proportionately more. The studies conclude that revenue recycling may be an interesting method offset the losses due to reduction in competitiveness and that in general; carbon taxes are indeed regressive unless subsidies are provided to the low carbon intensive companies. Moreover, carbon leakage is a said to be major a cause for concern [24] as companies may simply shift the manufacture of carbon intensive merchandise to countries without a carbon tax or with lax governmental regulations. The conclusions drawn from this discussion are presented below.

IV. Conclusion

The main conclusions of the above review can be summarized as follows:

1] From the discussion on industrial policies, it can be concluded that CCS is not yet a feasible option and could be made more cost effective by methods like Carbon Capture and Usage, where the stored carbon can then be used profitably.

Supplementary cementitious materials are already in use in the cement industry and further research has to be done to ascertain the exact amount of substitution that can be done in the cements without affecting their strength. If the strength is affected by the substitution, the substitution becomes redundant as more amounts of the same clinker will be needed to achieve the same strength.

2] Alternative fuels for the cement industry have already been explored exhaustively and not much potential for further CO2 reduction exists in that area. However, using alternative fuels may have other advantages to the environment.

3] The use of alternative cement clinkers with a different composition is one area that merits further research. If the cost analysis of different types of cements is done, they could begin replacing the OPC in the markets. For this, the market bankability of the Portland cement has to be taken into account. It would definitely help if experts and educationalists raised awareness about the benefits of using different types of cements and hence increase their reliability.

4] From the governmental policies it can be concluded that while Voluntary Approaches are flexible mechanisms, their effectiveness is very difficult to gauge and may require more standardization.

5] Carbon trading while having the advantage of giving flexibility to the companies who can decide which method they want to use to reduce their carbon emissions, has the disadvantage of not being transparent and being very easy to evade. Moreover, the method is expensive to execute and often complicated.

6] Carbon taxes on the other hand are simple enough in theory but may give rise to multiple complications. They are transparent and easy to implement. However, the imposition of carbon taxes raises quite a few problems. If set too high, they may affect the competitiveness of the industry and raise carbon leakage issues. If set too low, they lose their effectiveness as they do not provide sufficient incentive to the companies to reduce the emissions. If the companies are dissatisfied with the imposition of this tax, it may promote them to produce carbon in covert operations, which may cause more damage to the environment. Policies by the governments need to address the issue of competiveness directly through compensation mechanisms for the deserving companies. Moreover, most research finds that carbon taxes are regressive in nature.

7] In such a scenario, the design of a carbon tax is highly important and very few studies exist where the design of the carbon tax is given sufficient consideration. The design of the carbon tax should also be done taking into account different cements and technologies available. Moreover, most studies do not account for uncertainty and perform no risk analysis. This could be an important missing link as every factor affecting the carbon tax is subject to a lot of variability.

The promotion of a less carbon intensive cement industry requires that the government and the industries work together through a mixture of strategies and technologies based on the different domestic conditions of various countries.

in the unnumbered footnote on the first page.

References

[1] P. Kumar Mehta, "Reducing the environmental impact of cement", *Concrete International*, 2001.

[2] Alain Bilodeau and V. Mohan Malhotra, "High volume fly ash system: concrete solution for sustainable development", *ACI Materials Journal*, Title No. 97-M6, 2000.

[3] Joseph Davidovits, "Global warming impact on the cement and the aggregates industries", *World Resource Review*, vol. 6, no. 2, pp. 263-278, 1994.

[4] Zi-yue Chen and Pu-yan Nie, "Effect of carbon tax on social welfare: A case study in China", *Applies Energy* 183, pp. 1607-1615, 2016.

[5] Xianbing Liu, Yongbin Fan, and Can Wang, "An estimation of the effect of carbon pricing for CO_2 mitigation in China's cement industry", *Applied Energy*, 185, pp. 671-686, 2016.

[6] Walid Matar, Amro M. Elshurafa, "Striking a balance between profit and carbon dioxide emissions in the Saudi cement industry", *International Journal of Greenhouse Gas Control*, 61, pp. 111-123, 2017.
[7] Karen L. Scrivener, Vanderley M. John, and Ellis M. Gartner, "Eco-efficient cements: Potential, Economically viable solutions for a low CO_2, cement based industry", UNEP report, 2015.
[8] Cesar Valderrama, "Implementation of the best available techniques in cement manufacturing: a life cycle assessment study", *Journal of Cleaner Production*, 25, pp. 60-67, 2011.
[9] N. Muller and J. Harnisch, "A blueprint for a climate friendly cement industry", 2008.
[10] Yudiesky Cancio Diaz , Sofia Sánchez Berriela, Urs Heierlic, Aurélie R. Favierd,Inocencio R. Sánchez Machadoa, Karen L. Scrivenerd, José Fernando Martirena Hernándeza, and Guillaume Haber, "Lime stone Calcined clay cement as a low carbon solution to meet expanding cement demand in emerging economies", *Development Engineering*, 2, pp. 82-91, 2017.
[11] Benjamin C. McLellan, "Costs and carbon emissions for geopolymer pastes in comparison to ordinary Portland cement," *Journal of Cleaner Production*, 19, pp. 1080-1090, 2011.
[12] Mark Bediako, Charles Dela Adobor, Eric Opoku Amankwah, Kofi Nyako, and Charles Kwame Kankam, "Maximizing the sustainability of cement utilization in building projects through the use of greener materials", *Journal of Engineering*, vol. 2016, Article ID 1375493, 6 Pages, 2015.
[13] http://civilengineersforum.com/geopolymer-concrete-advantages-disadvantages/
[14] Shraddha Mishram and Dr. Nehal Anwar Siddiqui, "A review on environmental and health impacts of the cement manufacturing industry", *International Journal of Geology, Agriculture and Environmental Sciences*, vol. 2, 2014.
[15] World Steel Association, Worldsteel in figures, Brussels.
[16] Eugeniusz Mokrzycki and Alicja Uliasz- Bochenczyk, 2003, "Alternative fuels for the cement industry", *Applied Energy*, 74, pp. 95-100, 2015.
[17] Stephen Bygrave and Jane Ellis, "Policies to reduce greenhouse gas emissions in industry - Successful approaches and lessons learned: Workshop report", OECD and IEA report, 2003.
[18] Z. Eylem Gevrek and Ayse Uyduranonglu, "Public preferences to carbon tax attributes", *Ecological Economics*, 118, pp. 186-197, 2015.
[19] Boqiang Lin and Xuehui Li, "The effect of carbon tax on per capita CO_2 emissions", *Energy Policy*, 39, pp. 5137-5146, 2011.
[20] Tsai Chi Kuo, I-Hsuan Hong, and Sheng Chun Lin, "Do carbon taxes work? Analysis of governmental policies and enterprise strategies in equilibrium", *Journal of Cleaner Production*, 139, pp. 337-346, 2016.
[21] Xianbing Liu, Can Wang, Dongjie Niu, Sunhee Suk, and Cunkuan Bao, "An analysis of company choice policy preference to carbon tax policy in China", *Journal of Cleaner Production*, 103, pp. 393-400, 2014.
[22] Zhe Zhang, Aizhen Zhang, Daoping Wang, Aijun Li, and Huixan Song, "How to improve the performance of carbon tax in China", *Journal of Cleaner Production*, 142, pp. 2060-2072, 2016.
[23] Andrea Baranzini et al, "A future for carbon taxes", *Ecological Economics*, 32, pp. 395-412, 2016.
[24] Roger Sathre and Lief Gustavsson, "Effects of energy and carbon taxes on building material competitiveness", *Energy and Buildings*, 39, pp. 488-494, 2006.

National Conference on Recent Advances in Computer Science and IT (NCRACIT)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 1 | ISSN: 2456-3307

Cheater Detection Scheme for Dynamic Secret Key Generation

Sunil Dalal[1], Sangeeta Bhan[2], Susobhan Das[3]

[13]Department of Information Technology BGSB University, Rajouri, J&K, India
[2]Computer Science and Engineering, Delhi Technological University, Delhi, India
sunildalal57@gmail.com[1], sangeetabhan7@gmail.com[2], sdas1980@gmail.com[3]

ABSTRACT

Any secret sharing scheme is considered to be strong if the complexity of cryptanalysis is very high. In cryptography the brutal attack is the type of attack in which all the possible keys of the algorithm are tried for obtaining the secret. Various types of cryptanalytic methods for a given scheme are compared with complexity of the brute force attack. If the complexity of various cryptanalytic attacks is of the same order as that of brute force attack, then the system is said to be computationally strong. In our method of dynamic multi-secret sharing the security analysis mainly depends upon the application of hash function to calculate C matrix and modular arithmetic over primitive element. In this paper, two methods are proposed which remove the need of secret share distribution by the dealer using secure channel and to find out the cheater in the group.

Keywords: Secret Key, Cheater Detection, Cryptanalysis.

I. INTRODUCTION

In any kind of network [8] whether wired or wireless, communication reliability and security are two important issues, especially when networks are used in national security and safety critical applications. The secure and reliable way to send information on network [5] is to encrypt the information with the help of encryption algorithm and a secret key. So in this case we have to mainly protect the cryptographic [1] key in order to protect the encrypted data from disclosure as encryption algorithm is publically known due to limited in number and follow a definite pattern to encrypt data, so cryptanalysist can easily know about encryption algorithm [9]. The secure scheme keeps the key safe by putting it in single location but such a scheme is unreliable in case of a single misfortune occurs. To keep multiple copies of a key is also unsecure from theft point of view. Here we need more robust secure way to protect the key and this can be done by sharing of the secret key/ Password using various schemes [7].

There are situations, where a military troop has many missiles but all do not have the same launch code. To launch missiles in war by sharing the launch codes one can use single secret sharing method n-times. Alternatively Multiple-Secrets sharing method can be used. In, 2006 Li Bai [6] developed multiple-secret sharing scheme based on matrix projection. Li Bai method of multiple-secret sharing is also threshold method of secret sharing. But if military troop have to change the launch code at regular intervals to keep it safe without changing the secret shares values or new codes have to be assigned to new missiles without changing the secret share values then in such cases dynamic multi-secret

secret plan can be utilized. Many times this scheme is also not enough for giving the reliability such as we need to keep up a secret on-line. We can store the secret on server.

Therefore Lin-Yeh developed a dynamic multi secret sharing scheme that can be used, when we have to change the secret without need to change secret shares. Therefore this scheme is more capable to stop intruders. The cheaters are those who are shareholder in scheme or the dealer itself who is distributing wrong secret shares to users. In such cases, There should be a method for detect the cheater from good shareholders and also a method so that dealer not able to know about secret shares. Lin-Yeh scheme does not provide such kind of solution. We can apply here a verification process to detect the cheater and also the problem regarding dealer. Dynamic multi- secret sharing schemes includes secret shares selection by users themselves using modular arithmetic and verification can be done by calculating W matrix and publish it publically. Now the verification and cheater detection process take place.

II. RELATED WORK

The Shamir's Scheme [2] of Secret Sharing method consist of two stages, namely (a) Shares Generation Phase, (b) Secret Reconstruction Phase Steps followed in each of these stage as presented in are as follows:

A. Shares Generation Phase

1) Select the limit plot (k, n) for conveyance of secret and in view of estimation of k, compose a polynomial p(x) of degree k-1
2) It can be randomly chosen (k-1) degree i.e. polynomial according to mathematical equation, like $q(x) = (a_0 + a_1x + \dots + a_{k-1}x^{k-1}) \bmod P$ a_0 i.e. a_0 is a value to be used for secret key. Select a constant value for $a_1, a_2 \dots a_{k-1}$.
3) Find more valuable result for modular arithmetic to be applied on GF(P) ,hence P is greater than (S, n). Here S = secret value that will be shared.
4) Select n diverse estimations of x to get the distinctive estimations of y. The sets framed by gathering (y, x), are diverse offers for members. Appropriate these offers among the members.

B. Secret Reconstruction Phase

1) Gather k or more number of members share.
2) Utilizing Lagrange's introduction strategy, reproduce the secret. Secret from the interjection strategy is ascertained from this equation.

$$q(x) = \sum_{i=1}^{k}(D_i) \prod_{\substack{j=1 \\ j\neq i}}^{k} \frac{(x - x_j)}{(x_i - x_j)}$$

Here, put the value of x= 0 then formula is reduce t

$$q(0)=\sum_{i=1}^{k}(D_i) \prod_{\substack{j=1 \\ j\neq i}}^{k} \frac{(-x_j)}{(x_i - x_j)}$$

$$q(0)=\sum_{i=1}^{k}(D_i)\, L_i$$

q(0) provide the value of secret data. So, we can see that how one can distribute and reconstruct the secret using Shamir's secret sharing scheme.

In the same year i.e. in 1979, Blakely [3] developed a method of secret sharing based on hyper-plane. Two non-parallel lines intersect at exactly one point on a plane, three non-parallel planes intercepts exactly at one point. Analogous to this theory any n non-parallel planes intersect exactly at one point. Blakely proposed method of secret sharing by using the above said property of non-parallel planes. According to his algorithm, any secret to be shared can be encoded as single coordinate or all coordinates of the intersected point. His method of secret sharing was not as secure as to Shamir's method as any person (participant) who owns any one plane information can encode other plane because the point will also lie in the other plane. Therefore, it is not theoretically secure for information sharing.

III. PROPOSED METHOD

In proposed method of Dynamic multi-secret sharing scheme employs hash function [10] and modular arithmetic concept on primitive element. In first proposed method modular arithmetic operation are applied on primitive elements. Master secret shares xj are choosing by user. Then user calculate yj using equation ($y_j = g^{xj}$ mod p, where g is a primitive element) and send this value to the dealer. The dealer can checks that no two users choose same secret shares by comparing their respective y_j values he got from users. There is no secure channel is required by the dealer to distribute secret shares and users choose their master secret shares by themselves.

The proposed scheme called as "Cheater Detection scheme for Dynamic Secret Key Generation" consists of three stages namely (A) System initialization and construction stage (B) The Pseudo Secret Share Generation Stage (C) The Group Secret Reconstruction Stage The steps followed in each stage are as follow:

A. System Initialization and Construction Stage

In the framework initialization phase the dealer D is select the accompanying parameters.

1) P: A vast prime number can have any value depend on number of client and group secrets value, p -> (group secrets, n).

2) g: The primary component upon GF (p).

3) h(.): The safe one-way hash function that acknowledges contributions of the any dimension i.e. creates a settled length input.

4) IDj : With respect to the client the value assigned to Identifiers is Uj (user) for j = 1 to n. here n is the number of clients which are having the secret share value.

B. The Pseudo Secret Share Generation Stage

Suppose there is a Dealer with name D that can share group i.e. k secrets ($S_1, S_2, \dots, S_i$), for i= 1 to k i.e n clients.

1) Each user Uj (j=1, 2, 3⋯ n) select value of master secret share xj itself. Then it compute = g^{x_j} mod p, keeps xj secretly and send yj to the dealer by public channel so that dealer can ensure that yj' ≠ yj.

2) k number of secrets the D i.e. dealer has ($S_1,S_2,\dots S_k$). At that point he creates k is the no. of polynomial fi(x) of (i-1) degree, for i= 1 to k, comparing to the privileged insights takes after

f1(x) = S1 where f1 (0) = S1

f2(x) = S2 + d1x where f2 (0) = S2

.

.

.

fi(x) = Si + d1x + … +d(i-1) x(i-1) where fi(0) = Si

Where (S1,S2,…….,Si) are the insider facts themselves whose equation. that equation is polynomial to be computed and d1 to d(i-1) i.e. chosen arbitrarily by the merchant of any constants from GF (p).

3) To evaluate for i= 1 to k and j= 1 to n

Vij (i ϵ k, j ϵ n)= fi(IDj) mod p =

$$\begin{bmatrix} f_1(ID_1) & f_1(ID_2) & \cdots & f_1(ID_n) \\ f_2(ID_1) & f_2(ID_2) & \cdots & f_2(ID_n) \\ \vdots & \vdots & \ddots & \vdots \\ f_k(ID_1) & f_k(ID_2) & \cdots & f_k(ID_n) \end{bmatrix}$$

4) Each user compute C_ij(pseudo secret share) by using hash function and master secret share $\boldsymbol{x_j}$.

$C_{ij} = (h^i(x_j) \oplus x_j)$ mod p, where $h^i(x_j)$, indicates i progressive uses of h to x_j.

i.e. $C_{1j} = (h(x_j) \oplus x_j)$ modulo (p)

$C_{2j} = (h(h(x_j)) \oplus x_j)$ modulo (p)

.

.

$C_{kj} = (h^k(x_j) \oplus x_j)$ *modulo (p)*

And then sends C_{ij} to the dealer.

5) Dealer now compute

$\boldsymbol{R_{ij}} = \boldsymbol{V_{ij}} - \boldsymbol{C_{ij}}$ modulo (p).

$W_{ij} = h(h^i(x_j) \oplus x_j)$ *modulo (p).*

C. The Group Secret Reconstruction Stage

To reproduce m[th] bunch secret, called S_m out of (S_1,S_2 ,… ,S_k) i.e. S_m is belongs to (S_1,S_2 ,… ,S_k) at any rate m members out of n clients should helpfully play out

the accompanying strides with the gathering secret merge.

1) To compute the pseudo secret share of every user U_j, for j = 1 to m, (condition is m ≤ n) as follows:

C_{mj}= $(h^m (x_j) \oplus x_j)$ mod p and at that point it will post to the gathering secret merge over a protected channel.

The gathering secret combiner checks legitimacy of every member's pseudo secret share Cmj by the accompanying conditions:

$W_{mj} = h (h^m(x_j) \oplus x_j)$ mod p.

If any user is cheating at the time of reconstruction of secret by giving wrong secret share then we can check by analyzing hash function on its share scheme and match it with respective value in W_ij matrix and find that secret share given is right or wrong. This cheater detection technique is not available in Lin Yeh method.

2) After accepting all C_{mj}, for j = 1 to m, effectively gathering secret combiner remakes the mth assemble secret Sm takes after: [4]

$$S_m = \sum_{j=1}^{m}(C_{lj} + R_{lj}) \prod_{r=1,r\neq j}^{m} \frac{-ID_r}{ID_j - ID_r} \text{ MOD P.}$$

As framework R_{ij} is publically known in this way the gathering secret combiner get R_{mj} relating to the particular C_{mj} and ID_j of the individual clients from R_{ij} lattice (i.e. in the event that third secret is to be recreate and first, second and third clients are takes an interest in secret remaking then R_{3j} have values R_{31}, R_{32} and R_{33} taken from R_{ij} lattice known publically).

Thusly one can disperse insider facts and remaking of privileged insights should be possible by utilizing above depict plans. The plan can be effectively executed. In the event that another secret is to be included by merchant then he does not need to convey the ace secret shares among clients.

IV. RESULTS AND ANALYSIS

In Lin Yeh's scheme the dealer has to select master secret shares x_j where j=1 to n Where n is number of users) and master secrets Si i.e. i = 1 to k. where k is number of master secrets). Dealer distribute with the respective user's j master secret shares x_j over secure channel shown in Figure 1. Second, there is no method to verify the secret shares given by user at the time of secret reconstruction. Therefore, any user can give a wrong share to group secret combiner.

In proposed scheme the user's themselves select the master secret shares x_j (Therefor j= 1 to n. where n is number of users). Dealer has to select only master secrets Si (Therefore i = 1 to k. where k is number of master secrets). Each of the users calculates a value y_j using his master secret share x_j and a primitive element g by using equation $y_j = (g^{x_j})$ mod p and post it to the dealer over public channel, so that he can verify that no two y_j have same values shown in Figure 2. Therefore, there is no need of secure channel in initial distribution of master secret shares; moreover it is not possible for any other person to know about master secret shares. Second, there is a method to verify the secret shares given by user at the time of secret reconstruction. A $W_{(ij)}$ matrix is calculated by applying hash function on C_{ij} matrix. Therefore, in the event that a pernicious member U_j may post a false secret share C_{ij} which can be checked by the group secret combiner by applying hash function on the received C_{ij} values and verify it with the help of respective values in $W_{(ij)}$ matrix since he have the knowledge of $W_{(ij)}$ matrix.

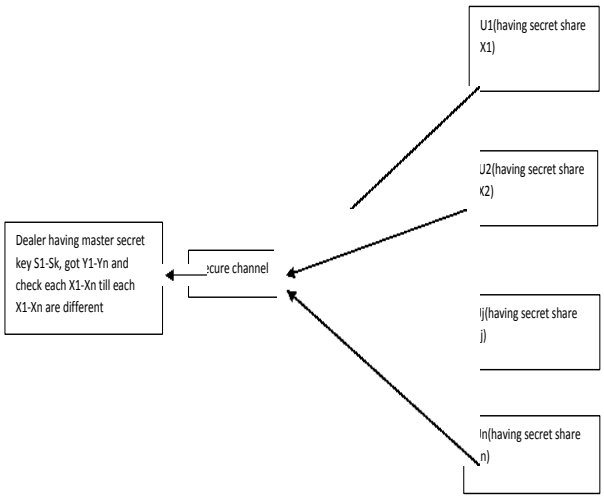


Figure 1.

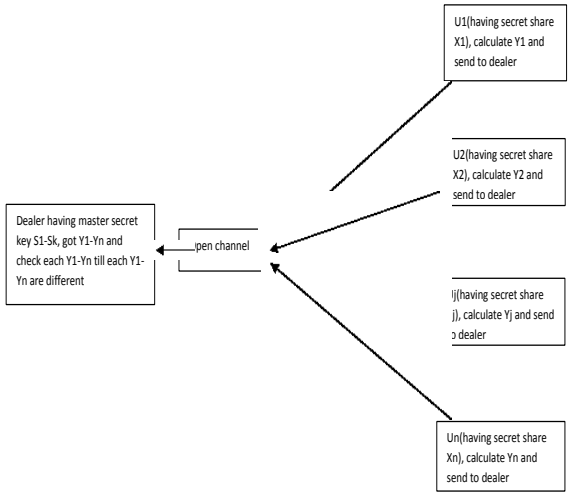


Figure 2.

The comparison of system performance among the two schemes using various parameters discussed. How our scheme takes equal time as taken by Lin-Yeh's scheme [4] but it provides more security with decrease in complexity level. Here, we introduce a method due to which even dealer cannot think of becoming a cheater. Various parameters are being discussed below in the Table 1.

Capability	Lin Yeh's scheme	Proposed Scheme
Need of security channel during secret share distribution	Yes	No
Need of security channel during sending pseudo secrets to dealer as well as group secret combiner.	Yes	Yes
Participant choose his secret share	No	Yes
Group secrets corresponding threshold value	Yes	Yes
The group secret combiner is able to check whether participant's pseudo secret share is true or not	No	Yes

V. CONCLUSION

The secret sharing is a critical technique to understand the disseminated data security utilizing information encryption. It is additionally a basic apparatus in multi-party setting. It has been connected in numerous applications, for example, secret mission information, and computerized money, amass marks et cetera. Over twenty years has gone since the main secret sharing convention has been designed. Various secret sharing plans were proposed from that point forward, which viably quickened the improvement of the hypothesis of data security and the broadly utilization of results of data security. Here we propose a dynamic secret sharing arrangement, in light of XOR operation, the unmanageability of discrete logarithm and hash capacities. The proposed plot has justifies then Lin-Yeh's plan. In proposed conspire each member picks the secret share by her/him. Is completely incomprehensible for merchant to wind up plainly a con artist and it needn't bother with a protected channel during initial distribution phase. The scheme has simple verifiable phase. So this scheme can be used in same practical cases such as secret sharing in electronics commerce, electronic govt. etc. The presented work can be extended to Insertion of authentication scheme and some more complicated cryptographic technique as elliptic cryptographic

curve techniques can be used at time of secret distribution process. The elliptic curve cryptography is a way to deal with open key cryptography in view of the arithmetical structure of elliptic curve over limited fields and has strong algebraic properties.

VI.REFERENCES

[1] Atul Kahate, Cryptography and network security, second edition, India: Tata McGraw-Hill.

[2] Shamir Adi, "How to share a secret", Communication of ACM, vol. 22(11), pp. 612- 613, November 1979.

[3] Blakely G., "Safeguarding cryptographic keying", In Proc. Of AFIPS, National computer conference, 1979.

[4] Lin Han-Yu and Yeh Yi-Shiung, "Dynamic Multi-Secret Sharing Scheme", Int. J. Contemp. Math. Sciences, Vol. 3, no.1, pp. 37-42, 2008.

[5] William Stallings, Cryptography and Network Security, third edition, India: Pearson Education.

[6] Bogdnav Dan, "How to securely perform computation on secret share data" ,Master's thesis, University of Tartu, 2007.

[7] Menezes, P. Oorschot, S. Vanstone, Handbook of applied cryptography, CRC Press.

[8] Hallberg, Networking a beginning guide, TMH.

[9] Diffie W., Hellman M., "New directions in cryptography", IEEE Transactions, on Information Theory, IT-22 (6), pp. 644-654, 1976.

[10] He J., Dawson E., "Multistage secret sharing based on one-way function", Electronics Letters, vol. 30 (19), pp. 1591-1592, 1994.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325070398

Design of Flying Wing UAV and Effect of Winglets on its Performance

Article · March 2018

CITATIONS
0

READS
19

5 authors, including:



Shabir Grover
Delhi Technological University
4 PUBLICATIONS **1** CITATION
SEE PROFILE



B B Arora
Delhi Technological University
33 PUBLICATIONS **53** CITATIONS
SEE PROFILE

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

Dates Extended for Full manuscript Submission at ICAPIE 2016, Delhi, India View project

All content following this page was uploaded by Shabir Grover on 10 May 2018.

The user has requested enhancement of the downloaded file.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

Design of Flying Wing UAV and Effect of Winglets on its Performance

Rishabh Dagur[1*], Vikrant Singh[1], Shabir Grover[1], Nikhil Sethi[1], Dr. B.B. Arora[2]

[1,2,3,4] *Undergraduate Student, Dept. of Mechanical Engineering, Delhi Technological University, India*
[5]*Professor, Dept. of Mechanical Engineering, Delhi Technological University, New Delhi, India*

Abstract- In the present work, a Flying wing body was designed that can be used for surveillance and reconnaissance. The UAV is highly portable and can be launched for a mission from almost anywhere. A Design space was mathematically plotted by defining parameters: cruising speed, stall speed, GTOW, aspect ratio etc. After defining design requirements, these values are used t o stipulate parameters in designing equations of UAV sizing and airfoil selection. During the design phase, it was understood that a UAV with increased endurance will be more helpful in practical use. Our study made us realize that winglet configuration can play a major role in reducing the drag of the body and, finally increasing the overall mission time of the UAV. Previously, not much work has been done on winglet design of Hand-Launched UAVs and the vorticity analysis. Therefore, we have taken up this as the main subject for our paper. The parameters for winglets were calculated using the standard values from the previous research. Both the flying wing geometries i.e., with and without winglets were modeled and solved in CFD code. Both the models were compared to see the effect of winglets in drag reduction which ultimately agreed to increase the lift and endurance efficiency of the flying wing body with the application of winglet. Also, a detailed analysis of vortex formation across the wingtip for both the scenarios was performed. Five Angle of Attacks were considered for both the cases.

Keywords: Unmanned Aerial Vehicles, Flying wing body, Winglets, Tailless-UAV, Computational Fluid Dynamics.

I. NOMENCLATURE

a	Coefficient of GTOW
b	Coefficient of empty weight equation
W_{to}	Maximum Take-Off weight
W_{pl}	Weight of Payload
W_c	Crew Weight
W_e	Empty Weight
W_f	Fuel Wight
V_s	Stall speed
V_{max}	Maximum speed
ROC_{max}	Maximum rate of climb
S_{TO}	Take-off run
ρ	Density of air
μ	Friction Coefficient
C_{lmax}	Maximum coefficient of lift
V_{max}	Maximum speed
P_{sl}	Engine/Motor Power
C_{d0}	Drag Coefficient at 0 Angle of Attack
ρ_0	Air Density
C_{DG}	Coefficient of Drag at Ground
C_{LR}	Lift Coefficient at Take-Off
K	Induce Drag Factor
η_p	Propeller Efficiency
AC	Aerodynamic Code
σ	Air density Ratio
B	Wing Span
S	Surface Area
C_l	Coefficient of lift
C_d	Coefficient of drag
C_d	Coefficient of drag
W/S	Wing Loading
W/P	Power Loading

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

II. INTRODUCTION

The aircraft industry is unconventionally breaching new horizons of research. One of the major sectors of development is Unmanned Aerial Systems (UAVs). UAVs are flying vehicles which can be remotely controlled from large distances (sometimes even from different continents). UAVs are divided on behalf of their size, endurance, number of motors, take-off and landing capabilities. A tailless flying wing UAV was designed under this research project. Fuselage geometry was maintained same as the airfoil along the wing. Our major aim in this research project is to make the aircraft endurance-efficient by decreasing drag on the body. After going through several research papers, the addition of winglets was taken up as the most elementary solution for our need. This also prevented any alterations in the design and added versatility. Winglets are vertically extruded wing tips incorporated to reduce lift induced-drag. Winglets are the part of the wing that which can be mathematically understood as the result of increased aspect ratio(given a constant area).

III. Literature Review

The concept winglet was coined by Dr. Whitcomb in 1976 [1] who at NASA research center and developed winglet technology which was published as "A Design Approach and Selected Wind-Tunnel Results at High Subsonic Speeds for Wingtip Mounted Winglets". According to him, winglet could be described as the small wing-like vertical structures which extend from the wingtip to reduce induced drag in comparison to other wing tip devices or extensions and claimed in his research that the winglet shows 20% reduction in induced drag when compared to tip extension and also improved the lift-to-drag ratio. In 1994, Aviation Partners Inc. (API) developed an advanced design of winglet namely blended winglet. Louis B. Gratzer from Seattle has the patent for blended winglet and intention of the winglet is to reduce the interference drag due to sharp edges as seen in the Whitcomb's winglet. Also, Gratzer has the patent for the invention of spiroid-tipped wing in April 7, 1992. Later, "wing grid" concept was developed by La Roche from Switzerland in 1996 and got the patent for his invention. The main purpose of all the above inventions was to decrease the vortex strength at the wingtip. Also, a remarkable work was done by Jamey and Jacob [22], they designed and optimized the Winglet using VLM methodology (Vortex Lattice Method) also known as Panel Method. The main advantage of this method is that wings are easy to design and it takes less Computational time to calculate Lift and Drag. But the disadvantage is that it solves the problem assuming the flow is inviscid, and the boundary layer is solved separately and both of the results are interpolated to find lift and drag, which raises a question on the validity of results as skin friction drag (which is dependent on viscous modeling) is neglected.

These are the major research work previously undertaken on winglets. Our main aim is to design, analyze and optimize the design using CFD(Computational Fluid Dynamics), which takes viscous flow into account.

Before undertaking this research, a background study was conducted which included an in-depth analysis on aerodynamics and wing design from works of John D. Anderson [3], Daniel P.Raymer [2], and by Mohammad Sadraey [6]. For the design of winglets, we referred Whitcomb's existing research on winglets. Our winglet design is majorly based on "Understanding Winglet technology by Fred George [7] and the research by, R. T. Jones and T. A. Lasinski [4].CFD by John D. Anderson [5] and ANSYS User guide [11]. As our main aim was to reduce drag by reducing vortex formation, a study was conducted on turbulent flows and vorticity, which was primarily done from books by S.K Som [10] and Cengel and Cimbala [8] and "Methods for Vortex Identification" by Vivianne Holm´en [9].

IV. Design Process of The Flying Wing

The design process is expounded to enfold all available design aspects. Algorithms were written in MATLAB to simulate the equations to secure the fittest solutions for the defined boundaries. A design space is produced to select the optimum value from the available space.
The whole process followed during the design can be defined as can be shown in following flow chart:

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

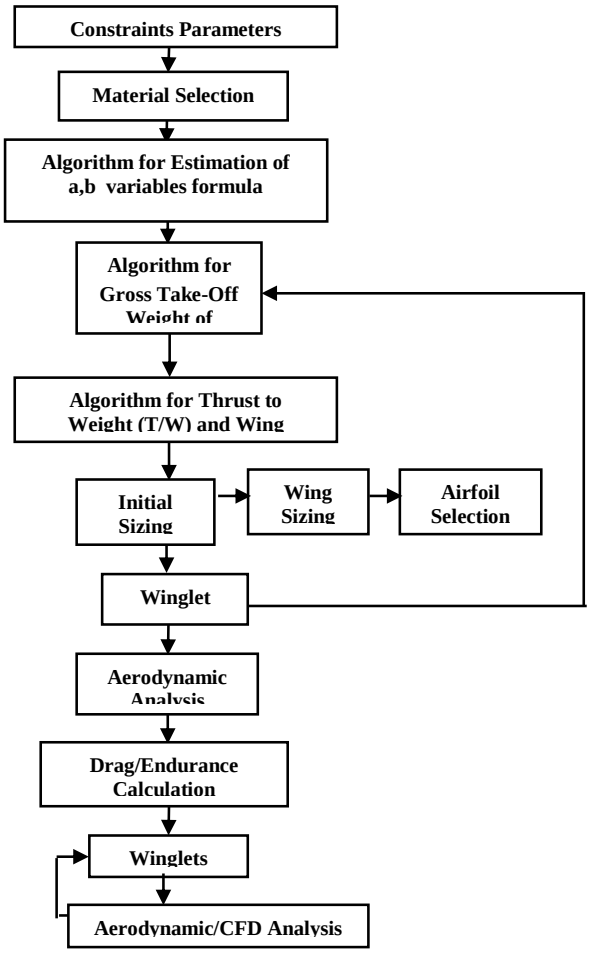


Fig-1: Flow Chart for Design Process

A. Constraint Parameters:

Design parameters were selected after an intensive market research of different UAVs of the same size and weight parameters in the market. Existing COTS airframes like Skywalker X8, Ebee, RVJet were considered while making the table.
From the below parameters defined, size and weight of the aircraft will play a critical in the whole designing process. Also, the maximum payload of the aircraft will be embodied by the gimbal and camera to give live-video feed.

Parameters	Objective	Weightage (Out of 5)
Wingspan	<2 meters	5
Cruise Speed	15 m/sec	2
Payload	>500 Gram	4
Endurance	30 minutes	4
Stall Speed	10 m/sec	2
GTOW	<5 kg	5
Take-Off Distance	Hand-Launched	4

Table-1: Design Parameters

B. Material Selection

Material selection was a crucial point where every available option in the market was tested to simulate real working conditions of the UAV missions (particularly 1. Take-off 2. Flight 3. Land/crash). Loads were applied to create destructive testing conditions and repairable attributes were judged by severity towards each impact.

Very Good	Good	Bad	Poor

S.No.	Material	Weighted	Strength /Weight	Reparability
1	Carbon Fiber	High		
2	Balsa Wood	Low		
3	EPO Foam	Low		
4	EPP Foam	Very Low		
5	Pink Foam	Low		

Table-2: Material Selection

From Table.2, EPP (Expanded Poly-propylene) was chosen as it is most promising for UAV missions of our class.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

C. Calculating "a" And "b" Variables for Empty Weight EQ.

Initially, the weight of flying wing is not known exactly, so the GTOW must be estimated which relies on the other aircrafts with a similar configuration and mission available in the market. Thus, historical data is the major source of information for the calculations in this step. Following is the collective data from available flying wings in the market.

UAV	We	Wp	Wb	GTOW	We/Wo	Wp/Wo	Wb/Wo
Ebee	0.77	0.2	0.646	1.616	0.4764	0.12376	0.399
Zephyr2	2.97	0.424	1.569	4.963	0.598	0.0854	0.0316
Eflite Opterra	4.187	2.5	1.12	7.8	0.5363	0.3202	0.1434
RVJet	4.85	2.2	2.4	9.45	0.51322	0.2328	0.2539
Skywalker X8	6.04	3.19	2.7	11.93	0.5062	0.2673	0.2263
Viper X10	6.61	3.2	2.5	12.31	0.5369	0.2599	0.203

Fig-2: Comparison on Weighting Chart

The two main equations that underline calculations in this step are as follows:

$$W_{to} = \frac{W_{pl}+W_c}{1-\left(\frac{W_f}{W_{to}}\right)-\left(\frac{W_e}{W_{to}}\right)} \quad \ldots\ldots (1)$$

$$\frac{W_e}{W_{to}} = aW_{to} + b \quad \ldots\ldots.(2)$$

Equation 1 shows the net summation of the GTOW and fuel weight of the aircraft. Also, as our UAV is electrically powered thus,

$W_c = 0;\ W_f = W_b$

∴ Equation (1) will become,

$$W_p = W_{to} - W_b - W_e \quad \ldots\ldots (3)$$

The second equation depicts the relationship between the empty weight fraction and GTOW which was scaled to form a linear programming problem in MATLAB.

The different values of weight were grabbed from already available market UAVs. Eqn. (3) and (2) are used to solve for "a" and "b" as there are 2 unknowns and 2 equation that is a system of equation in 2 variables. Different available UAVs give different vales of "a" and "b". The most optimum design space for the values of "a" and "b" was elected on the basis of mission requirements from the from below plotted graph in MATLAB.

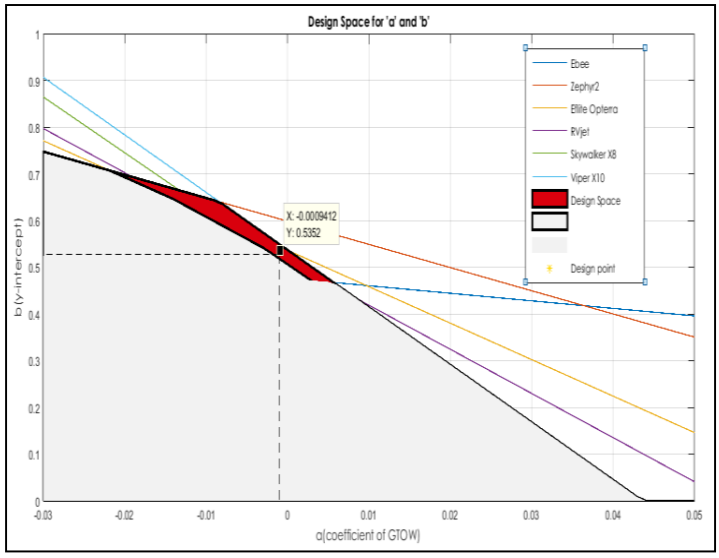


Fig-3: Design Space for 'a' and 'b'

D. Determining Gross Take-Off Weight:

To find the value of GTOW, MATLAB code was written to find the solution of the two governing equations, i.e. modified Eqn. 1 i.e. Eqn. 3 and Eqn. 2

The estimated weight of the battery and payload was defined from market research and constraint parameters analysis.

A MATLAB program was written and value of GTOW was determined by solving the quadratic equation.

Weight	Estimations(lbs.)
W_p	1.5
W_b	2.2

Table-3: Weight Estimation

By analytically solving the quadratic equation in MATLAB, 2 values for GTOW were obtained. Out of the given values obtained one is 485 pounds which is not feasible and impossible value for our requirement, so **8.09 pounds** was selected. Thus the GTOW is taken as 8.5 lb. for a safe design. Thus we have tolerance as ((8.5-8.093)*100/8.093) = **5.02 % Tolerance**

Variables	Approximation
a	9.412 e-04
b	0.5352

Table-4: 'a' and 'b' Values

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

E. Thrust to Weight Ratio and Wing Loading

This phase is solely dependent upon the aircraft performance requirements and employs flight mechanics theories. Hence, the technique is an **analytical approach.**

The aircraft performance requirements deployed to size the aircraft in this step are:

- Stall speed (V_s)
- Maximum speed (V_{max})
- Maximum rate of climb (ROC_{max})
- Take-off run (S_{to})
- Ceiling

The equations and graphs have the power loading as the independent variable and wing loading as the dependent variable. The equations used are:

- Stall speed:

$$\frac{W}{S} = \frac{1}{2}\rho {V_s}^2 W C l_{max}$$

- Maximum speed:

$$\left(\frac{W}{P_{sl}}\right) V_{max} = \frac{\eta}{\frac{\frac{1}{2}C_{do}\rho {V_{max}}^3}{\frac{W}{S}} + \frac{2K}{\rho V_{max}}\left(\frac{W}{S}\right)}$$

- Take-off Run :

$$\left(\frac{W}{P}\right)_{S_{TO}} = \frac{1 - \exp\left(0.6\rho g C_{D_G} S_{TO} \frac{1}{W/S}\right)}{\mu - \left(\mu + \frac{C_{D_G}}{C_{L_R}}\right)\left[\exp\left(0.6\rho g C_{D_G} S_{TO} \frac{1}{W/S}\right)\right]} \frac{\eta_p}{V_{TO}}$$

- Rate of Climb

$$\left(\frac{W}{P}\right) ROC = \frac{ROC}{\frac{ROC}{\eta} + \sqrt{\frac{2}{\rho\sqrt{\frac{3C_d}{K}}}\left(\frac{W}{S}\right)}\left(\frac{1.155}{\left(\frac{L}{D}\right)\max \eta_p}\right)}$$

- Ceiling

$$\left(\frac{W}{P_{sl}}\right) = \frac{\sigma A C}{\frac{ROC}{\eta} + \sqrt{\frac{2}{\rho\sqrt{\frac{3Cd}{K}}}\left(\frac{W}{S}\right)}\left(\frac{1.155}{\left(\frac{L}{D}\right)\max \eta_p}\right)}$$

The code is written on MATLAB for the above defined equations and iterations are run. Out of the graph, optimum values are selected according to constraints defined. Each different point will give changing values of wing and power loading which would eventually change the wing area and power required for it. On running the code the following graph is obtained:

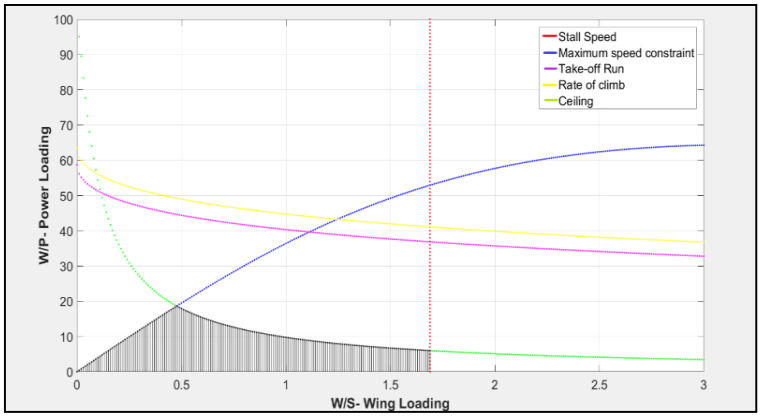


Fig-4: Power Loading vs Wing Loading

The design point which gives the lowest wing loading with the lowest power loading(power loading given the higher priority f0r lowest selection) selected as,

X: Wing loading=**W/S=1.69 lb/ft2.**

Y: Power loading=**W/P=11.78 lb/hp.**

Now, the wing reference area and power can be easily found by dividing GTOW by X and Y:

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

S.no	Attributes	Equation	Estimation
1	Wing area	W/(W/S)	8.5/1.69 = 5.029586 ft^2 or 0.46726m^2
2.	Motor Power	W/(W/P)	8.5/11.78= 0.721Hp or 538.285 W

Table-5: Reference Area and Power Required

F. Sizing

After defining the wing's surface area, the geometry of the wing is defined and parameters like wing sweep and taper ratio are referenced from already justified models as the calculations are judged on the real-life command of the aircraft with the reflection with air.

Since these are mere ratios, the overall weight and area will influence the design.

In general, a tailless fixed wing has the following similar geometry characteristics:

- Very low taper ratio
- Very low aspect ratio
- High sweep angle at the quarter chord
- Winglets to provide directional stability
- Negative washout/Tip twist

The UAV which is taken as reference is the Skywalker X8 due to its flying capabilities and its stability-control when compared with the other aircrafts.

For Skywalker X8, these values are extracted for its definition.

S.no	Attributes	Values
1.	Aspect ratio	5.618
2.	Taper ratio	0.4
3.	Leading edge sweep	27°

Table-6: UAV Skywalker X8Attributes Study

Using the following formulas, Aspect Ratio (AR):

Defined as span of wing B, to mean aerodynamic cord C,

$$AR = \frac{B}{C}$$

Wing area can be defined as span of wing multiply with mean aerodynamic cord.

Wing Area(S):

$$S = B.C$$

From, the above two equations aspect ratio can be mathematically formulated as,

$$AR = (\frac{B}{B}).\frac{B}{C} = \frac{B^2}{S}$$

Wingspan= 5.584 ft.
Mean Aerodynamic Chord= 0.9003 ft.
Root chord = 1.246 ft.
Tip chord= 0.5548 ft.

To compare the parent UAV, Skywalker X8 and the designed flying wing, the below spreadsheet is defined as,

	Skywalker X8	Flying wing
Wing Loading	1.72862192	1.69
Thrust loading	4.451492537	11.78
Wing Area	6.9014513	5.029585799
Power required	2.68	0.721561969
Wing span	6.54166	5.584494464
Root Chord	1.46	1.246375067
Tip Chord	0.65	0.554893009
MAC	1.055	0.900634038
Aspect Ratio	6.200625592	6.200625592
Taper ratio	0.445205479	0.445205479

Fig-5: Design Comparison with Skywalker X8

G. Aerofoil Selection

Airfoil for an aircraft involves the consideration of multiple interrelated parameters and the integration of trade-offs between design requirements. The airfoil selected for the UAV is **'s5010'**.

The primary flight requirements and trade-offs that need to be made defined as:

1. A good C_l/C_d ratio. To increase loiter time.
2. Since it does not have a tail and lacks stability in general, a trade-off will have to be done between **performance and excellent stall characteristics.** A gentle transition after the stall angle to enable easy recovery.
    a. The aerofoil with the lowest (~0) pitching moment coefficient.
    b. Aerofoil with net positive pitching moment.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

3. An aerofoil with a proper ideal and design lift coefficient (same reason as 1st)

Low speed, high lift – (t/c) max=15-18%
High speed, low lift- (t/c) max=9-12%

The selection process is purely graphical and deals with the comparison of some shortlisted airfoils on certain graphs like:
C_l Vs. alpha
C_m Vs. alpha
C_d Vs. C_l
C_l/C_d Vs. alpha
XFLR5 is the open source software package used to analyze the above graphs.

Taking into consideration of self-stabilizing (reflexed) some of the airfoils shortlisted were:

- E186
- MH-45
- MH-60
- S5020
- S5010

The graphs acquired from XFLR5 are used to input in the spreadsheet.
The required ideal lift coefficient and the gross maximum lift coefficient is calculated from following equations:

$$Cl_{ideal} = \frac{Wempty}{\left(\frac{1}{2}\right)\rho V^2 S}$$

$$Cl_{gross\ max} = \frac{Wmax}{\left(\frac{1}{2}\right)\rho V^2 S}$$

The stall characteristics are given a score out of 5.

Similarly, the C_l/C_d max and the slope is also compared as given in the design requirements.

On the basis of comparison and trade-offs, the airfoil **S5010** which provides excellent stall characteristics and sufficient lift is selected.

GTOW	Wing Area	Criuse Speed	Stall Speed	CL_cruise	CL_ideal	CL_max	Cl/Cd	Cd min	Stall Characterstics
10	7.815	49.21	32.8	0.44033	0.4892	1.13275			
	t/c								
S5020	8.4				0.4861	1.1587	74.6	0.0081	4
S5010	9.82				0.4961	1.19334	73.59	0.078	5
MH-45	9.85				0.41	1.215	70.089	0.00752	3
MH-60	9.85				0.41	1.189	1.189	0.0076	2
E-186	10.27				0.187	0.892	55.7	0.0094	1

Fig-6: Chart for Airfoil Selection.

V. Winglet Design

Winglet is the wing tip extension that is used to reduce the vortex formation at the wingtip. These vortices are caused by pressure differences above and below the wing, causing sideways motion of the air. These vortices cause downwash of the flow stream about the wing tip, causing Induced Drag.
The winglets are designed so as to reduce the vortex by the stopping sideways flow, thus reducing the induced drag.
There are various winglets that are being used in aircraft industries like blended sharklets, Fenced winglets, Spiroid winglets and raked wing tips.
We chose Sharklets as they are suited for low speed flying aircraft, and are easy to manufacture.
Parameters affecting winglet Design: Sweep angle, Cant Angle, taper ratio, Span length (height). Experimentation indicated that the winglet's trailing edge should be positioned near the wing's trailing edge for maximum effectiveness [7]. Whitcomb found the elliptical load distribution could be preserved if the winglet had approximately a 0.3 taper ratio and a side force loading about the same as the wing loading [1]. Giving the winglet a slight cant angle or dihedral also improves its aerodynamics. Whitcomb [1] found that this reduced the interference at the junction of the winglet and wingtip at transonic speeds and that it pushed the tip vortex outboard, thus further reducing nearby vortex intensity. The optimum cant angle for Whitcomb's experiment was 15 degrees.
Usually, height of winglet was taken around 10-20%. We took 11.1% as suggested by the NASA reports [1].

Specifications:

- Air foil: S5010 (same as that of wing)
- Tip chord of the wing (winglet root chord): 0.155 m
- Height of winglet: 0.085m
- Taper Ratio: .3 (Standard value for winglets on most commercial aircrafts)
- Tip chord of winglet: 0.076m

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

VI. CFD Analysis of Body without Winglet

The CFD analysis of the wing was done on 5 angles of attacks: 0, 2, 4, 6, 8 degrees. The software used is ANSYS and the code is Fluent.

A. Geometry:

The CAD geometry was made in DS Solidworks. It was then imported in ANSYS Design Modular in
".step" file. First, as the flying wing Body is symmetrical, therefore it was divided into half which helped in restraining the elements to half thereby reducing simulation run time. The enclosure was then made of dimension 3m*1m*1m. Then using the Boolean tool the wing body was subtracted from the enclosure.
This process was done separately for each angle of attack.

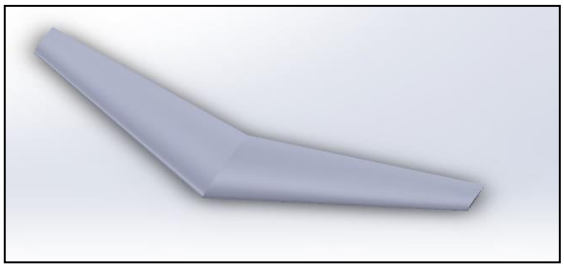


Fig-7: Cad Model of Wing without Winglets

B. Meshing:

The Meshing was done in ICEM. Initially, the Named Selections where created which included an inlet, outlet, wing face, wing end, ambient wall etc. Then meshing attributes like minimum element sizing, maximum face sizing, inflation layers were edited to suit our case. To ensure low skewness elements, prism elements are extruded from the wing surface in the first step. After this, the rest of the domain is populated by the generic tetrahedral elements. The triangle height on surfaces must vary gradually to ensure good prism characteristics. Subsequent prism layers should show a constant geometric rate of increase in their height. A good transition at the interface between the prismatic layers and the tetrahedral region is a must to ensure a mesh satisfying the numerical requirements in terms of cell-size deviation. Thus, it is imperative to find the right balance of growth rate, first aspect ratio and total number of layers to avert numerical diffusion.

The choice of the turbulence model used depends on the importance given to near wall effects or how the flow behaves near the wall. For this, it is it is important to know about the concept of wall Y^+ .

If our aim is to resolve the effects close to the wall i.e., in the viscous sub layer then the size of the mesh size should be small and dense enough near the wall so that nearly all the effects are captured efficiently. However, in some cases, when the wall effects can be ignored, a semi-empirical formula may be incorporated to bridge between the viscosity affected region and fully turbulent region. Hence, the mesh needn't be dense or small near the wall i.e., coarse mesh would work.

Based on the characteristic scales of our geometry the Reynolds number is determined for our model such that:

$$R_e = \frac{\rho UL}{\mu} = \frac{1.225 \times 15 \times 0.368}{1.789 \times 10^{-5}} = 3.7797 \times 10^5$$

Where ρ and μ are the fluid density and viscosity respectively, U is the free stream velocity, and L is the characteristic length (e.g. pipe diameter, body length, Chord length etc.)
The definition of the y+ value is such that:

$$y+ = \frac{rho \cdot U_\tau \cdot \Delta y_1}{\mu} \rightarrow \Delta y_1 = \frac{y+ \cdot \mu}{\rho \cdot U_\tau}$$

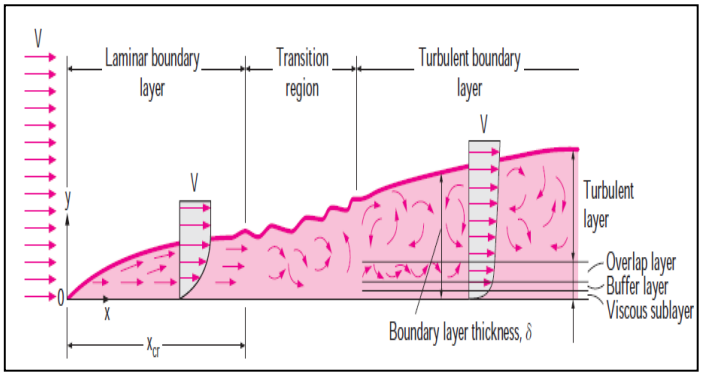


Fig-8: The boundary layer developed over flat plate (Çengel, Cimbala[8])

The known values are plugged into the above equation to estimate Δy_1(First cell height).
Given that walls are the main source of turbulence and mean vorticity, the near-wall modeling significantly determines the consistency of numerical solutions. This is because it is in the near-wall region that the solution variables have large gradients, and the momentum and other scalar transports occur most vigorously. Thus, for the successful prediction of wall bounded flows, the near wall region needs to be accurately modelled.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

The k-ε models, the LES, and the RMS model are primarily used for core turbulent flows (i.e., regions far from walls). However, the Spalart-Allmaras and k-w models are distinct in the sense that they can be applied throughout the boundary layer, provided that the mesh resolution is sufficient close to the wall.

Sizing	
Use Advanced Size Function	On: Proximity and Curvature
Relevance Center	Fine
Initial Size Seed	Active Assembly
Smoothing	High
Transition	Slow
Span Angle Center	Fine
Curvature Normal Angle	Default (18.0 °)
Num Cells Across Gap	5
Min Size	1.e-003 m
Proximity Min Size	6.e-004 m
Max Face Size	0.10 m
Max Size	0.10 m
Growth Rate	1.150
Minimum Edge Length	0.147310 m

Fig-9: Mesh Sizing

In our case, where the low-Reynolds-number effects are dominant throughout the flow domain, the placement of the first node in our near-wall inflation mesh becomes very important. The wall function approach would have been inadequate and thus, invalid. But, if the case involves high-Reynolds-number flows where the viscosity-affected near-wall region does not need to be resolved, the wall function approach is more suitable as it saves computational resources.

It is desirable to resolve the boundary layer all the way to the wall using a finer mesh for an accurate prediction of the separation point and flow separation.(Example- In case of determining the drag or lift forces experienced by wings)

The wall Y+, apart from being dependent on the first layer thickness, also depends upon the local fluid velocity. That aspect cannot be neglected. And that is why, it is important that the Y+ values are checked regularly as part of the post-processing procedure so as to ensure that it lies in the valid range for our flow physics and turbulence model selection.

Consequently, for a y+ value approximately equal to 1 with 10-15 cells within the boundary layer thickness renders any wall models useless as we are already resolving the flow all the way to the wall. A y+ value between 30 and 300 would work sufficiently if we want to work on a coarse mesh with wall functions to capture the velocity profile close to wall. In the laminar sub-layer region or Y+ < 5 inertial forces are less dominating and the flow is primarily laminar, this is known as the low-Re region. Low-Re turbulent models such the SST model as mentioned earlier aim to resolve this area and thus require an appropriate mesh refinement to be efficacious. This is paramount for flows with a changing pressure gradient and where we expect to see flow separation.

Definition		
Suppressed	No	
Boundary Scoping Method	Named Selections	
Boundary	wing face	wing end
Inflation Option	Total Thickness	
Number of Layers	10	
Growth Rate	1.2	
Maximum Thickness	2.e-002 m	1.e-002 m
Inflation Algorithm	Pre	

After grasping the concept of wall Y+, inflation layer was generated at wing face and wing end separately. In all the 5 AOA cases, it was found that grid independence was reached around 2.1 million cells approximately. The specification of meshing attributes were the same for all 5 AOA and are as follows:

Fig-10: Inflation Settings

The elements in the mesh were of Generic tetrahedral structure.

AOA	Elements	Nodes
0º	2,725,678	4,677,429
4º	2,567,772	4,785,634
8º	2,215,345	4,213,457

Table-7: Mesh Stats for the given 3 AOA's

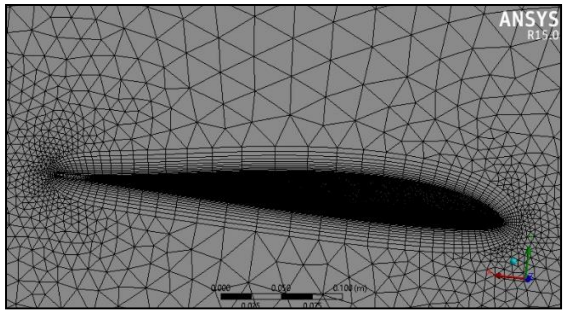


Fig.-11: Mesh Preview without Winglet

C. Solution:

The Simulation was then performed using FLUENT code. In boundary condition, the velocity was kept 15m/s and outlet gauge pressure was set as 0 Pa.

Selecting the most appropriate wall function is directly dependent on the level of refinement of the mesh in proximity to wall. This would affect the accuracy of the solution.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

Standard wall function such as ε-based models work on Y+ >30 which implies that our boundary layer mesh lies completely within the log-law region of the boundary layer and we can work with coarse mesh. But in practice, this may be difficult to achieve because of the varying geometrical and velocity scales associated with the model.
Scalable wall function offers a novel alternative. This wall function displaces the mesh to a Y+ ~ 11.225 for any level of refinement and this assists in avoiding the erroneous modelling of the laminar sub-layer and buffer region.
Performance on meshes of intermediate resolution can also be improved if we use Enhanced wall treatment for ε-based models on refined low-Re grids. However the use of non-equilibrium (enhanced) wall treatment for low-Re modelling of the turbulent boundary layer is generally not recommended and a suitable ω-based formulation, such as the SST model gives reliable results.

The use of a low-Re model is recommended advisable if our aim is to maximize the accuracy of the predicted boundary layer velocity profile or thermal profile, or if the developing boundary layer will tend to separate (due to a varying pressure gradient and not because of discontinuities in the geometry). Low-Re models are also required for drag calculations or determining pressure-drop accurately. The quality of the numerical results obtained will depend on the overall resolution of the boundary layer as discussed before. Thus, the low-Re SST model is the most suitable choice since the use of the k-omega formulation for the boundary layer region helps in resolving up till the viscous sub-layer and no extra wall functions are introduced. In the free-stream, the SST model switches to a k-epsilon formulation which is better suited to free shear flows.
Turbulence model used was SST $\kappa-\omega$ turbulence model. The standard $\kappa-\omega$ model in **ANSYS FLUENT** is based on the Wilcox $\kappa-\omega$ model [25]. This Wilcox Model subsumes modifications for low-Reynolds-number effects, compressibility, and shear flow spreading [24]. The Wilcox model estimates free shear flow spreading rates that are in close agreement with measurements for far wakes, mixing layers, and plane, radial, and round jets, and is applicable to both wall-bounded flows and free shear flows [25].
The Shear-Stress Transport (SST) $\kappa-\omega$ model was originally developed by Menter, to effectively blend the robust and accurate formulation of the $\kappa-\omega$ model in the near-wall region with the free-stream independence of the $\kappa-\varepsilon$ model in the far field. To achieve this, the $\kappa-\varepsilon$ model is converted into a $\kappa-\omega$ formulation. The SST $\kappa-\omega$ model is basically a blend of the standard k-w and the transformed k-e model. The SST $\kappa-\omega$ model is similar to the standard $\kappa-\omega$ model, but includes the following refinements [11]:

- This blending function is incorporated/multiplied into the k-e and k-w equations and is so designed that near the wall, the standard k-w model is active while as we move away from the surface, the effect of $\kappa-\omega$ fades away and the k-e model becomes prominent.
- To account for the turbulent shear stress transport, a modified definition of turbulent viscosity is used.

The turbulence kinetic energy, κ and the specific dissipation rate, ω are obtained from the following transport equations:

$$\frac{\partial}{\partial t}(\rho k)+\frac{\partial}{\partial x_i}(\rho k u_i)=\frac{\partial}{\partial x_j}\left(\Gamma_k\frac{\partial k}{\partial x_j}\right)+\tilde{G}_k-Y_k+S_k$$

And,

$$\frac{\partial}{\partial t}(\rho\omega)+\frac{\partial}{\partial x_i}(\rho\omega u_i)=\frac{\partial}{\partial x_j}\left(\Gamma_\omega\frac{\partial \omega}{\partial x_j}\right)+G_\omega-Y_\omega+D_\omega+S_\omega$$

SST model produces good results, certainly much better when dealing with partially separated flows.
Solver used was "Pressure-Based" and the Pressure-Velocity coupling algorithm used was "Coupled". The pressure-based solver allows you to solve your flow problem in either a segregated or coupled manner. Using the coupled approach offers some advantages over the non-coupled or segregated approach. The coupled scheme obtains a robust and efficient single phase implementation for steady-state flows, with superior performance compared to the segregated solution schemes.
The coupled algorithm solves the momentum and pressure-based continuity equations together. The pressure-based coupled algorithm obtains a more robust and efficient single phase implementation for steady-state flows.
The first half of total iteration ran on 1st order upwind scheme to stabilize the continuity, the second of iteration ran on 2nd order upwind Scheme so as to achieve a higher order of accuracy.
The convergence criteria was kept to none.

D. Post Processing/Contours:
Following were the drag and lift results for 5 AOA.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

Table-8 : Lift & Drag values for different AOA's

AOA	Lift	Drag
0º	4.383N	0.508N
2º	8.56N	0.645N
4º	11.19N	0.748N
6º	16.02N	0.992N
8º	19.703N	1.24N

Other than examining lift and drag, vortex formation on the wing ends was also studied. ANSYS CFD Post Processing has following methods to define, identify and visualize vortices:

1. Helicity
2. Lambda 2 criterion
3. Q-Criterion
4. Swirl strength

Helicity is referred to as the dot product of vorticity and velocity vector.

$$H = (\nabla \times \vec{V}) \cdot \vec{V}$$

The rest 3 methods are based on Velocity based tensor:
The velocity gradient tensor D can be written as
$D_{ij} = \partial_{ui}/\partial_{xj}$.
As this is a second order tensor it can be broken into a symmetric and a skew-symmetric part
$D_{ij} = S_{ij} + \Omega_{ij}$
Where $S_{ij} = 1/2(\partial_{ui}/\partial_{xj} + \partial_{uj}/\partial_{xi})$
And $\Omega_{ij} = 1/2(\partial_{ui}/\partial_{xj} - \partial_{uj}/\partial_{xi})$.
S_{ij} is known as the rate-of-strain tensor, and Ω_{ij} is the vorticity tensor [9]. The characteristic equation for ∇u is given by
$\lambda^3 + P\lambda^2 + Q\lambda + R = 0$ [9] ,
Where P, Q, and R are the three invariants of the velocity gradient tensor [9]. Decomposition into symmetric and anti-symmetric parts these invariants can be expressed as follows.
$P = -tr(D)$
$Q = 1/2(tr(D)^2 - tr(D^2)) = 1/2||\Omega||^2 - ||S||^2$
$R = -det(D)$

The **Q-Criterion** methodology defines a vortex as a "connected fluid region with a positive second invariant of ∇u" i.e. Q > 0 [9]. In this a secondary condition on the pressure is added, requiring it to be lower than ambient pressure in the vortex [9]. Looking at the definition of the second invariant we can see that Q represents the local balance between shear strain rate and vorticity magnitude, defining vortices as areas where the vorticity magnitude is greater than the magnitude of rate-of-strain.

The **λ2-criterion** looks for a pressure minimum but removes the effects of unsteady straining and viscosity by discarding these terms. On evaluating the gradient of the Navier-Stokes equations, we get:

$a_{ij} = -1/\rho(p_{ij} + \nu u_{i,jkk})$

Where a_{ij} is the acceleration gradient and p_{ij} is symmetric part. Decomposing the acceleration gradient into symmetric and anti-symmetric parts we get the vorticity transport equation as the anti-symmetric part and the symmetric part:

$DS_{ij}/Dt - \nu S_{ij,kk} + \Omega_{ik}\Omega_{kj} + S_{ik}S_{kj} = -\rho/p_{ij}$

The first two terms on the left-hand side represent unsteady irrotational straining and viscous effects respectively [9]. Therefore, only $S^2 + \Omega^2$ is considered to determine if there is a local pressure minimum that entails a vortex [9]. A vortex is defined as "a connected region with two negative eigenvalues of $S^2 + \Omega^2$" [9]. Since $S^2 + \Omega^2$ is symmetric it has real eigenvalues only, and by ordering the eigenvalues $\lambda 1 \leq \lambda 2 \leq \lambda 3$ the definition becomes equivalent to requiring that $\lambda 2 < 0$ [9]. Generally visualized as isosurfaces for different values of $-\lambda^2$. In planar flows, the three conditions that were described above are equivalent.
Q-criterion is widely a used method for vortex determination method and was also suggested when asked on various CFD Forums. It provides insight into the vorticity aligned with the fluid stream.
Following contours were observed for the 3 AOAs out of 5:

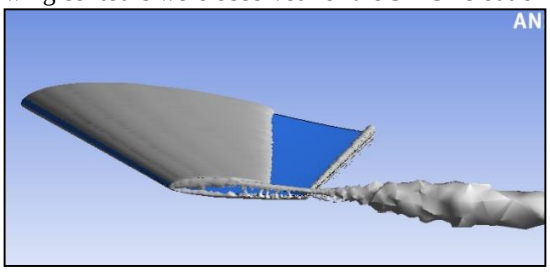


Fig-12: for 0º AOA, without Winglets

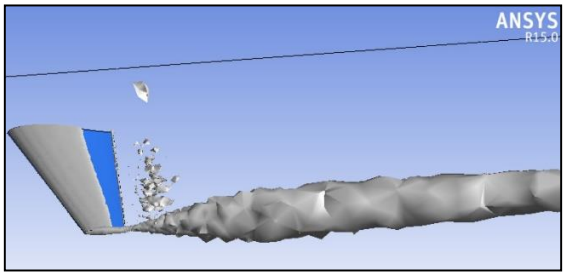


Fig-13: for 4º AOA, without Winglets

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

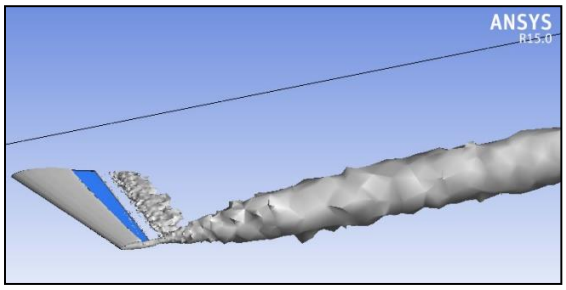


Fig.-14: for 8º AOA, without Winglets

Q-criterion values for body with winglets for 0.0004 level:

AOA	Q-Criterion values
0º	11.12 s^{-2}
2º	14.93 s^{-2}
4º	16.88 s^{-2}
6º	17.23 s^{-2}
8º	18.07 s^{-2}

Table-9: Q-Criterion Values for the 5 AOA's

Hence, we see, as the AOA increases, Vortex formation increases dramatically, thus increasing the drag.

VII. CFD Analysis of Body with Winglets

A. Geometry:

Same steps were followed as on body without winglets. The CAD geometry was made in SolidWorks. It was then imported in ANSYS in ".step" file. First, as the flying wing Body is symmetrical therefore it was divided into half then Enclosure was made of dimension 2m*1m*1m. Then using the Boolean tool the wing body was subtracted from the enclosure. This process was done separately for each angle of attack.

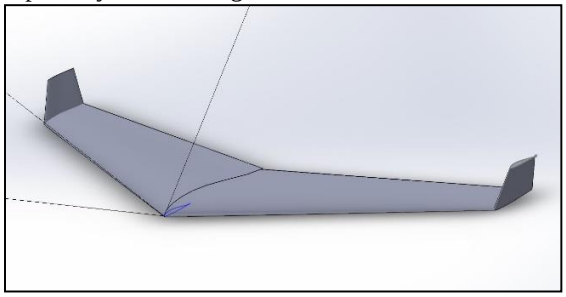


Fig-15: CAD Model with Winglet

B. Meshing:

The Meshing was done in ICEM. First Named Selections where created which included an inlet, outlet, wing face, winglet end, ambient surroundings etc. Then general face sizing was done. After that inflation was given at wing face and winglet end separately. In all the 5 AOA cases it was found that grid independence was reached around 4.2 million cells approximately. The specification of meshing attributes were the same for all 5 AOA and are as follows:

Sizing	
Use Advanced Size Function	On: Curvature
Relevance Center	Fine
Initial Size Seed	Active Assembly
Smoothing	Medium
Transition	Fast
Span Angle Center	Fine
Curvature Normal Angle	Default (18.0 °)
Min Size	5.e-004 m
Max Face Size	0.10 m
Max Size	0.10 m
Growth Rate	1.20
Minimum Edge Length	4.165e-004 m

Fig-16: Mesh Sizing

Suppressed	No		
Type	Element Size		
Element Size	5.e-003 m		
Behavior	Soft		
Curvature Normal Angle	Default		
Growth Rate	1.20	1.15	
Local Min Size	Default (5.e-004 m)		
Boundary Scoping Method		Named Selections	
Boundary		wing face	winglet end
Inflation Option		Total Thickness	
Number of Layers		12	10
Maximum Thickness		2.e-002 m	1.e-002 m
Inflation Algorithm		Pre	

Fig-17: Inflation Settings

The elements in the mesh were the generic tetrahedral structure.

AOA	Elements	Nodes
0º	4,243,789	8,815,647
4º	4,726,504	9,115,945
8º	4,587,697	8,678,765

Table-10: Mesh Stats

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

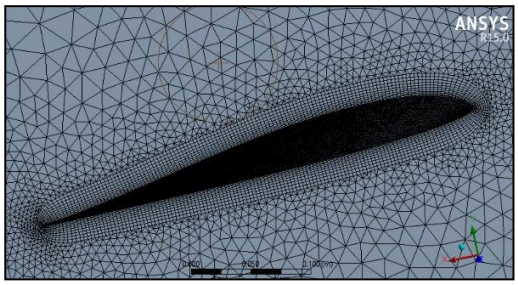


Fig-18: Mesh Preview, Wing with Winglet

C. Solution

The Simulation was then performed using the FLUENT code. In boundary condition, the velocity was kept 15m/s and outlet gauge pressure was set as 0 Pa. The same turbulence model SST k-w model was used. Solver used was Pressure-Based and the Pressure-Velocity coupling algorithm used was "Coupled".

The first half of total iteration ran on 1st Order Upwind scheme, the second of iteration ran on 2nd order Upwind Scheme so as to achieve a higher order of accuracy.

The convergence criteria was again kept to none.

D. Post processing/Contours:

In this case too we examined the vortex formation and following were the results:

AO	Lift	Drag
0º	4.86N	0.523N
2º	8.95N	0.638N
4º	12.57N	0.727N
6º	16.13N	0.956N
8º	20.64N	1.22N

Table-11: Value of Lift and Drag at few AOA's

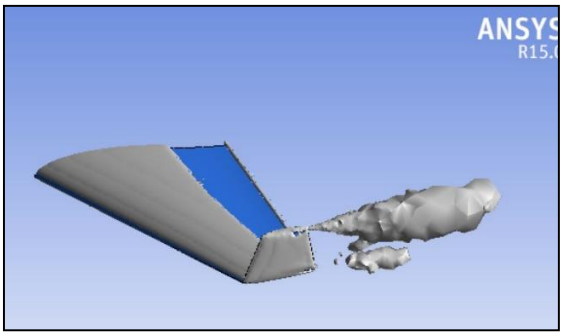


Fig-19: at 0º AOA

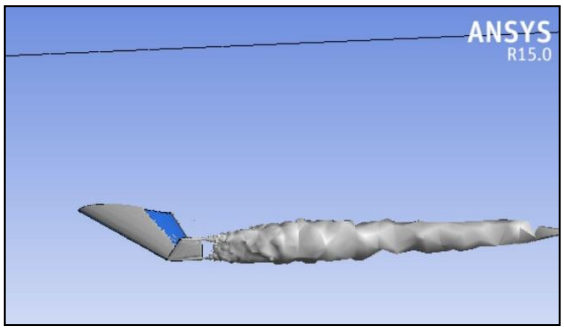


Fig-20: at 4º AOA

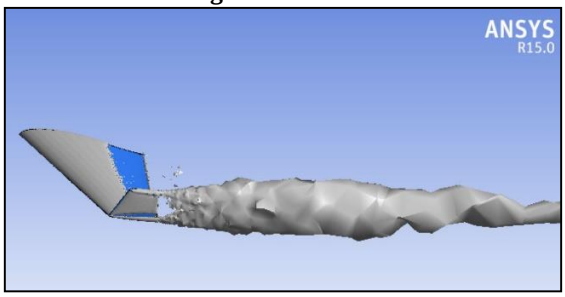


Fig-21: at 8º AOA

Q-criterion values for body with winglets for 0.0004 level:

AOA	Q-criterion Values
0º	11.41 s^{-2}
2º	12.34 s^{-2}
4º	13.78 s^{-2}
6º	14.67 s^{-2}
8º	16.18 s^{-2}

Table-12: Value of lift and drag at various AOA

VIII. Comparison of Results

Comparing both we can clearly see that with incorporation of winglet there was a slight increase in Drag at 0 AOA (2.95%) as compared to the model without winglets at same AOA. After further analyzing this anomaly we came to the conclusion that this is because at 0 AOA, viscous component are more prevalent than the pressure component of the drag.

But as the AOA increases the pressure component becomes more prevalent i.e. the vortex formation effect has a greater effect on the aerodynamics of the wing.
At 4º AOA there is a drag reduction of 2.1% and at 8º the reduction is 1.64%.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 8, Issue 3, March 2018)

We can see that there is slight decrease in Drag reduction of UAV as angle of attack increases to 8º because when seen overall the drag tends to increase when as AOA increases. We understand it well from equation of drag force:

$$F_d = \frac{1}{2}\rho V_s^2 C_d Area$$

Therefore there is a net decrease in drag reduction at 8º. With incorporation of winglets, it was observed that vortex formation was reduced due to blocking of sideways flow arising due to pressure difference. Also, it was found that the new reduced vortex of low intensity was being formed at winglet end, as shown in contours, of the wing with winglet's post processing, pushing the vortex formation away from the planform area. Both of these above statements imply that by the incorporation of winglets we can reduce the influence of vortex formation, which reduce the downwash, decreasing the induced drag.

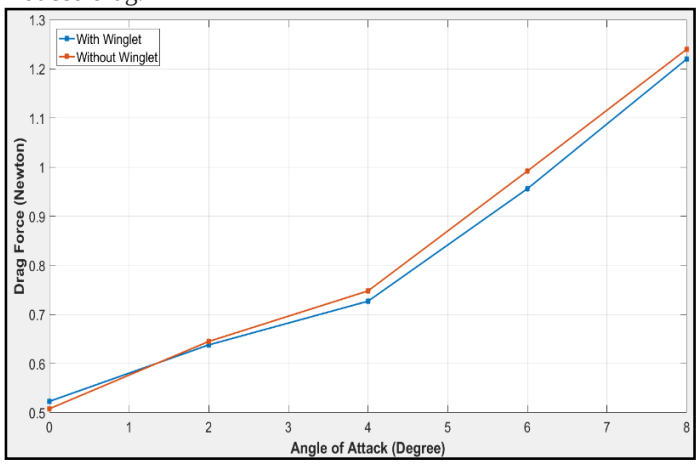


Fig-22 Comparison of Drag Forces for Wing, with Winglet and without Winglet

IX. Acknowledgements

The success of the completion of this project is strongly attributed to Dr. B.B. Arora under whose guidance the project has been taken. His knowledge and constant motivation went a long way which shaped the outcome in a beautiful manner. We would also like to thank our families and peers who acted as pillars of support throughout the project's completion.

X. References

[1] R. T. Whitcomb, "A Design Approach and Selected Wing-Tunnel Result at High Subsonic Speed for Wing-Tip Mounted Winglets," NASA TN D-8260, 1976.
[2] *Daniel P. Raymer*, "Aircraft Design: A Conceptual Approach" (AIAA Education Series) American Institute of Aeronautics and Astronautics, 1989 - Technology & Engineering - 729 pages.
[3] John D. Anderson, "Fundamentals of Aerodynamics", Boston: McGraw-Hill, 2001, Pages-892.
[4] R. T. Jones, T. A. Lasinski, "Effect of Winglets on the Induced Drag of Ideal Wing Shapes", NASA TM-81230, 1980.
[5] John D. Anderson, "CFD", McGraw-Hill Education, 01-Feb-1995 - Science - 547 pages.
[6] Mohammad H. Sadraey, "Aircraft Design: A Systems Engineering Approach", John Wiley & Sons, 28-Nov-2012 - Technology & Engineering pages-778.
[7] Fred George, "Understanding Winglet Technology", pp. 1-7.
[8] Yunus A. Çengel, John M. Cimbala, "Fluid Mechanics: Fundamentals and Applications", 2006, pages-994.
[9] Vivianne Holm´en, "Methods for Vortex Identification", November 21st 2012.
[10] S. K. Som, G. Biswas, "Introduction to Fluid Mechanics and Fluid Machines", 3rd Edition, Tata McGraw-Hill, 2007 – Fluid mechanics - 712 pages.
[11] ANSYS FLUENT USER'S GUIDE.
[12] B.B.Arora, Manoj Kumar, Subhashis Maji, "Study of Inlet conditions on Diffuser Performance," International Journal of Theoretical and applied Mechanics, Vol.5, No.2. (2010) pp. 201-221, ISSN 0973-6085.
[13] B.B.Arora, Manoj Kumar, Subhashis Maji, "Analysis of flow separation in wide angle annular diffusers," International Journal of Applied Engineering Research, Vol.5, No.20. (2010) pp. 3419-3428, ISSN 0973-4562.
[14] B.B. Arora, R.K. Sharma, A. Gogoi, Vipin and J.N. Rai, "CFD Analysis of a Free Power Turbine for an Auxiliary Power Unit," International Journal of Theoretical and Applied Mechanics, Volume 5 Number 2 (2010) pp. 223-232.
[15] Laurent Dumas, "Optimization and Computational Fluid Dynamics", pp. 191-215, Springer Berlin Heidelberg Germany, 2008.
[16] Marco Lanfrit, "Best practice guidelines for handling Automotive External Aerodynamics with FLUENT", Fluent Deutschland GmbH Germany.

[17] Chainani. A, Perera. N, "CFD Investigation of Airflow on a Model Radio Control Race Car", WCE 2008 Vol. II, July 2 - 4, 2008, London, U.K
[18] Yeung W. W. H., Parkinson G.V., "Analysis and Modeling of Pressure Recovery for Separated Reattaching Flows", ASME Journal of Fluids Engineering, Vol. 126, No. 3, pp. 355–361, 2004
[19] Wolf-Heinrich Hucho, "Aerodynamics of Road Vehicles", SAE International, 1998.
[20] Joseph Katz, "Race Car Aerodynamics: Designing for Speed", 1995.
[21] Shabir Grover, "Analysis on Drag Reduction of Bluff Body using Dimples", International Journal of Advanced Production and Industrial Engineering IJAPIE, Sp.Issue, pp. 4-11, 2017.
[22] Jacob *Weierman*, Jamey D. Jacob, "Winglet design and Optimization for UAVs" , AIAA-2010-4224
[23] Sesha Prakash and Unmanned Aerial Systems Technology Sector, Lockheed Martin Corporation, United States of America.
[24] Junjanna G.C, Dr. N Kulasekharan, Dr. H.R Purushotham, "Performance Improvement Of A Louver-Finned Automobile Radiator Using Conjugate Thermal CFD Analysis", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181 - Vol. 1 Issue 8, October – 2012.
[25] Wilcox, D. C., 1998, Turbulence Modeling for CFD, DCW Industries Inc., La Canada, California.

View publication stats

J. Inst. Eng. India Ser. B
https://doi.org/10.1007/s40031-018-0330-z

ORIGINAL CONTRIBUTION

Designing of Supplementary Controller for STATCOM for Damping of Oscillations in Power System

Vipin Jain[1] · **Narendra Kumar**[2]

Received: 18 February 2015 / Accepted: 17 April 2018
© The Institution of Engineers (India) 2018

Abstract In FACTS devices supplementary signals are widely used to enhance damping and mitigation of sub-synchronous resonance in power system. Subsynchronous resonance occurs due to series capacitor in the power systems. High value of series capacitive reactance may destabilize low frequency mode which is more dangerous. There are four critical values of series compensation (X_C) against which rotor oscillations may be very high in IEEE first benchmark model. In this paper modeling of STATCOM with IEEE first benchmark model is presented. Then a supplementary signal is developed which is capable to make the system stable for all critical values of series compensation. The eigenvalues are presented against all four critical values of series compensation.

Keywords STATCOM · Supplementary signals · Power system stability · IEEE first benchmark model

Introduction

The primary function of the static synchronous compensator (STATCOM) is to control the transmission line voltage at the point of connection but when supplementary signal is used as a feedback signal then it can enhance the damping of the power system. Feedback supplementary signal are widely used to enhance the damping of generator and Mitigate Subsynchronous Resonance (SSR) in series compensated power system [1]. Various new techniques and algorithms are developed to design damping controllers in recent years. An intelligent algorithm named Imperialism Competitive Algorithm (ICA) which is inspired by sociopolitical process of Imperialism is developed in Ref. [2] to design VSC-HVDC controller. It is also used to find the optimum parameters of conventional PSS. A damping controller based on Honey Bee Mating Optimization (HBMO) is designed for STATCOM and compared with Genetic Algorithm (GA) based controller in [3]. The effectiveness and robustness of the proposed controller are demonstrated through nonlinear time-domain simulation and some performance indices studying the damp low frequency oscillations under different operating conditions are presented. Evaluation results show that the HBMO based tuned damping controller achieved good performance for a wide range of operating conditions and is superior to the controller designed using GA technique. A scheme of parameter tuning of damping controller based on Artificial Intelligence (AI) is reported in [4]. In this study a new fuzzy logic controller is investigate for STATCOM and successfully tested on IEEE second benchmark model. Same controller is tested for Induction Generator also. All these controllers proposed in the literature are of academic interest and have not been adopted in practice; therefore a practical solution is needed. In STATCOM type '1' or type '2' converters can be used. In type '1' converters both active and reactive power in transmission line can be controlled. In type '1' converters DC battery is provided in parallel with capacitor and both K_{cs} and 'α' can be controlled. Magnitude of E_s, magnitude of STATCOM bus voltage and voltage across DC side capacitor is controlled through K_{cs} therefore reactive power in the system is controlled through K_{cs}. Active power of STATCOM and

✉ Vipin Jain
vipinjain71@rediffmail.com

[1] Electrical Engineering Department, Bharat Institute of Technology, Meerut, Uttar Pradesh 250103, India

[2] Electrical Engineering Department, Delhi Technological University, Delhi 110042, India

Published online: 26 May 2018

angle of STATCOM bus voltage is controlled through 'α'. DC battery can supply power in watts only but by controlling the voltage angle at STATCOM bus, active power in transmission line can be controlled. In type '2' converters, only 'α' can be controlled and 'K_{cs}' is kept fixed. 'α' alone can control Magnitude and angle of E_s, voltage across DC side capacitor, active power consumed by the STATCOM and reactive power consumed (or supplied) by the STATCOM. In type '2' converters, STATCOM cannot supply active power; it consumes active power to compensate the transformer losses and switching losses. Type '2' converters can controls magnitude of voltage at STATCOM bus and hence reactive power in the transmission line. For 'p' pulse converter $K_{cs} = (p/6)(\sqrt{6}/\pi)$ [5–7]. For 12 pulse converter $K_{cs} = 2\ (\sqrt{6}/\pi)$. 'α' is the angle difference between V_1 (STATCOM bus voltage) and E_s. 'α' is kept very small (maximum and minimum value of 'α' is kept ± 1°). Low value of 'α' can control desired reactive power and voltage of the system [8–12]. For low value of 'α', Reactive power supplied by STATCOM have linear relationship with 'α'. As the 'α' varies, V_{dc} (Voltage across DC side capacitor) varies, then magnitude of 'E_s' varies. ($|E_s| = K_{cs}V_{dc}$) which controls reactive power supplied (or absorbed) by the STATCOM [13–17]. In this paper type '2' converter is used. Study system considered in this paper is IEEE First Benchmark Model (FBM) [18]. Though any value of series compensation (X_c) can generate SSR problem but there are four critical values of X_C against which a particular mode of FBM may be highly unstable. These values are X_C = 0.185, 0.284, 0.379, 0.473 pu [19–21]. Small signal analysis is carried out against all four critical values of X_C for the damping controller designed in this paper. Results of Transient response is shown against X_C = 0.473 pu, which is the highest value of X_C among the all four aforesaid values.

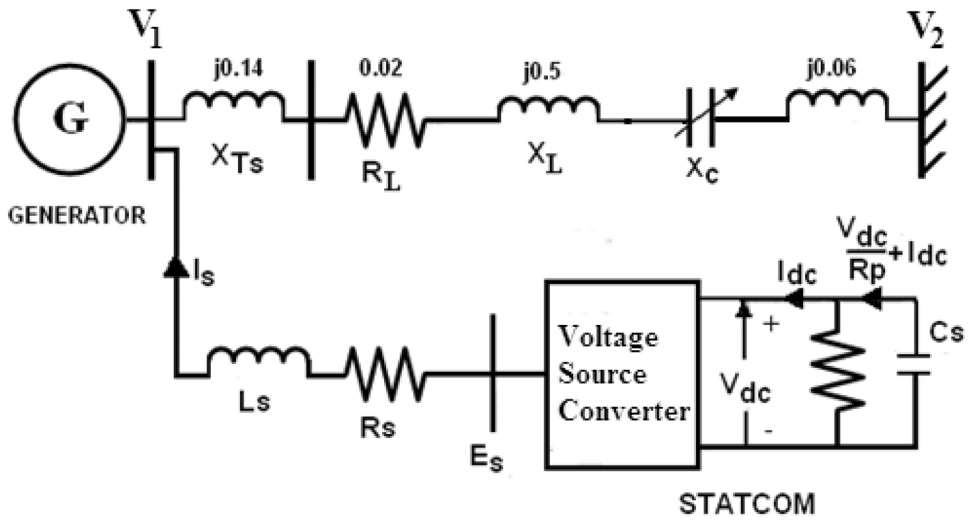


Fig. 1 STATCOM with IEEE first benchmark model

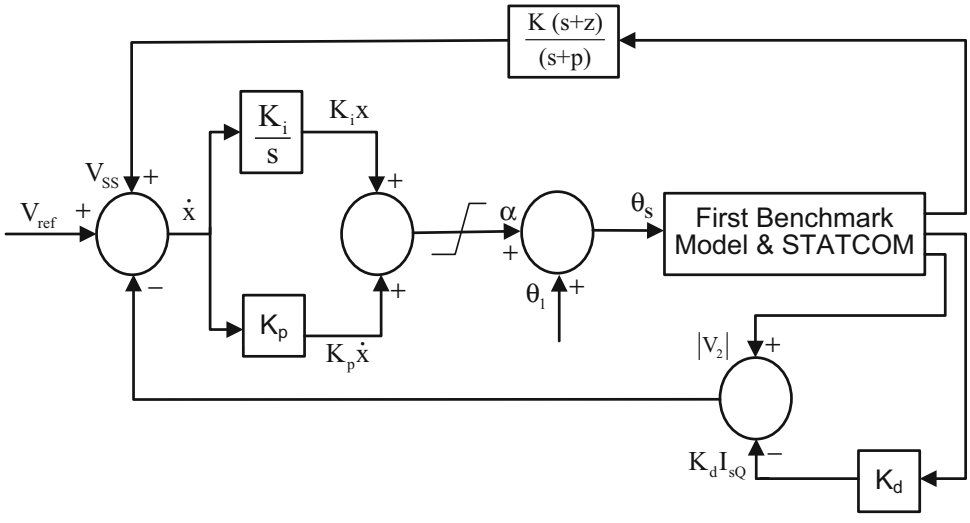


Fig. 2 Block diagram of the complete system with PI controller and supplementary signal

Study System

IEEE first benchmark model has twenty differential equations. Its generator equivalent circuit has six differential equations, mechanical system has twelve differential equations and Transmission network (i.e. series capacitor) has two differential equations. STATCOM is incorporated at Generator bus as shown in Fig. 1. X_{Ts} is the sending end transformer leakage reactance. STATCOM has three differential equations. Two differential equation of STATCOM are due to current through STATCOM (I_{sD} and I_{sQ}) and one differential equation is due to V_{dc}. Generator and turbine system have six masses. The block diagram of the system is shown in Fig. 2. In Fig. 2, output of PI controller is 'α', which is input to the STATCOM. Angle of bus E_s is $\theta_s = \theta_1 + \alpha$. θ_1 is angle of V_1. In case of disturbance θ_1 varies but as an input to θ_s it remains constant due to Phase Locked Loop (PLL). PLL synchronizes GTO pulses to the system voltage and provides a reference angle to the measurement system. In lineaisation $\Delta\theta_s = \Delta\theta_1 + \Delta\alpha$, but $\Delta\theta_1 = 0$. Hence $\Delta\theta_s = \Delta\alpha$. K_d represents droop characteristics of STATCOM. V_{SS} represents supplementary signal. V_{SS} must be zero in steady state. If $|V_1| > V_{ref}$ then $\dot{x}$ will be negative and 'α' will be positive (as the K_p and K_i are negative) and it may touch to α_{max}. In this case STATCOM is considered to be inductive mode. In capacitive mode 'α' will be negative. Generator equivalent circuit is shown in Figs. 3 and 4, where I_{1d}, I_{1q}, I_{2q} are the damper winding currents, which are zero in steady state. I_f is field winding current. X_l is leakage reactance of generator.

Differential Equations of the System

All the differential equations are written here. Equations 1–18 and 21 and 22 represent FBM dynamics. Equations 19 and 20 are output of transmission line system. Equations 23, 24 and 26 represent STATCOM dynamics; Eq. 27 shows PI controller dynamics. Equation 28 is

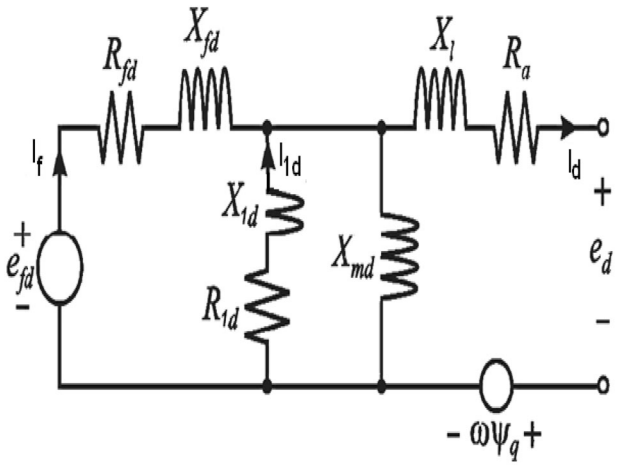


Fig. 3 Generator d-axis equivalent circuit

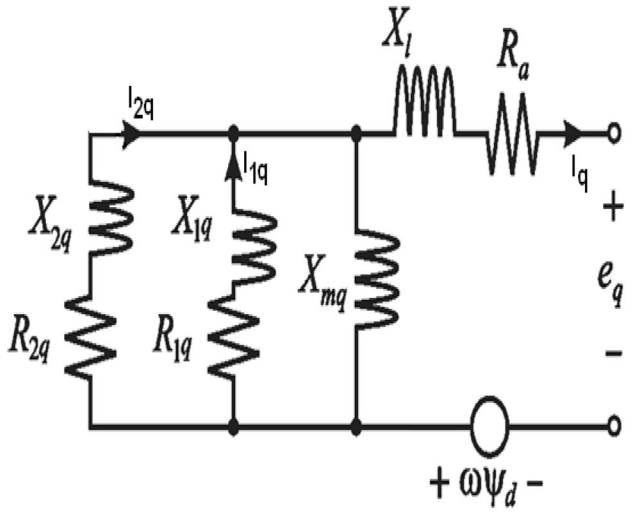


Fig. 4 Generator q-axis equivalent circuit

output of PI controller. Output of STATCOM are currents I_{sD}, I_{sQ}, $\dot{I}_{sD}$ and $\dot{I}_{sQ}$, which is input to transmission line system. Input to STATCOM is STATCOM bus voltage (or Generator terminal voltage) V_{1D} and V_{1Q}. Equations 1–6 are obtained from Figs. 3 and 4 and expressed in d–q axes.

Transmission line and STATCOM equations are expressed in D–Q axes. D–Q (network axes) and d–q axes (machine axes) both are rotating at synchronous speed in steady state period. In transient period speed of D–Q axes remain same but speed of d–q axes oscillate. Relationship between D–Q axes and d–q axes is shown in Fig. 5.

Generator Equations

$$\begin{aligned} e_d &= -R_a I_d - \frac{(X_l + X_{md})}{\omega_0}\dot{I}_d + \frac{X_{md}}{\omega_0}\dot{I}_f + \frac{X_{md}}{\omega_0}\dot{I}_{1d} - \omega\psi_q \\ \psi_q &= -(X_l + X_{mq})I_q + X_{mq}I_{1q} + X_{mq}I_{2q} \\ &\quad - \frac{(X_l + X_{md})}{\omega_0}\dot{I}_d + \frac{X_{md}}{\omega_0}\dot{I}_f + \frac{X_{md}}{\omega_0}\dot{I}_{1d} \\ &= R_a I_d - \omega(X_l + X_{mq})I_q + \omega X_{mq} I_{1q} + \omega X_{mq} I_{2q} + e_d \end{aligned} \quad (1)$$

$$\begin{aligned} e_q &= -R_a I_q - \frac{(X_l + X_{mq})}{\omega_0}\dot{I}_q + \frac{X_{mq}}{\omega_0}\dot{I}_{1q} + \frac{X_{mq}}{\omega_0}\dot{I}_{2q} + \omega\psi_d \\ \psi_d &= -(X_l + X_{md})I_d + X_{md}I_f + X_{md}I_{1d} \\ &\quad \times \frac{-(X_l + X_{mq})}{\omega_0}\dot{I}_q + \frac{X_{mq}}{\omega_0}\dot{I}_{1q} + \frac{X_{mq}}{\omega_0}\dot{I}_{2q} \\ &= \omega(X_l + X_{md})I_d + R_a I_q - \omega X_{md} I_f - \omega X_{md} I_{1d} + e_q \end{aligned} \quad (2)$$

$$\frac{-X_{md}}{\omega_0}\dot{I}_d + \frac{X_{fd} + X_{md}}{\omega_0}\dot{I}_f + \frac{X_{md}}{\omega_0}\dot{I}_{1d} = -R_{fd}I_f + e_{fd} \quad (3)$$

$$\frac{-X_{md}}{\omega_0}\dot{I}_d + \frac{X_{md}}{\omega_0}\dot{I}_f + \frac{X_{1d} + X_{md}}{\omega_0}\dot{I}_{1q} = -R_{1q}I_{1q} \quad (4)$$

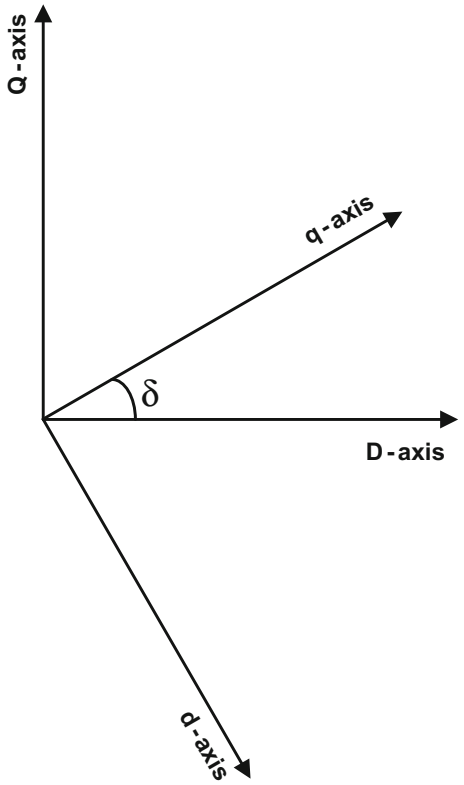


Fig. 5 D–Q reference frame (rotating at synchronous speed) and d–q machine axis frame

$$\frac{-X_{mq}}{\omega_0}\dot{I}_q + \frac{X_{1q} + X_{mq}}{\omega_0}\dot{I}_{1q} + \frac{X_{mq}}{\omega_0}\dot{I}_{2q} = -R_{1q} I_{1q} \tag{5}$$

$$\frac{-X_{mq}}{\omega_0}\dot{I}_q + \frac{X_{mq}}{\omega_0}\dot{I}_{1q} + \frac{X_{2q} + X_{mq}}{\omega_0}\dot{I}_{2q} = -R_{2q} I_{2q} \tag{6}$$

Mechanical System Equations

Exciter system equations are:

$$2H_E\dot{\omega}_E = K_{GE}(\delta - \delta_E) - D_E(\omega_E - \omega_S) \tag{7}$$

$$\frac{1}{\omega_0}\dot{\delta}_E = \omega_E - \omega_S \tag{8}$$

Generator rotor equations are:

$$2H_G\dot{\omega} = K_{BG}(\delta_B - \delta) - T_e - K_{GE}(\delta - \delta_E) - D_G(\omega - \omega_S) \tag{9}$$

$$\frac{1}{\omega_0}\dot{\delta} = \omega - \omega_S \tag{10}$$

where T_e is electrical torque generated by the generator

$$T_e = \psi_d I_q - \psi_q I_d$$

$$T_e = -(X_l + X_{md})I_d I_q + X_{md} I_f I_q + X_{md} I_{1d} I_q + (X_l + X_{mq})I_q I_d - X_{mq} I_{1q} I_d - X_{mq} I_{2q} I_d$$

LP-B turbine equations are:

$$2H_B\dot{\omega}_B = T_{LPB} + K_{AB}(\delta_A - \delta_B) - K_{BG}(\delta_B - \delta) - D_B(\omega_B - \omega_S) \tag{11}$$

$$\frac{1}{\omega_0}\dot{\delta}_B = \omega_B - \omega_S \tag{12}$$

LP-A turbine equations are:

$$2H_A\dot{\omega}_A = T_{LPA} + K_{IA}(\delta_I - \delta_A) - K_{AB}(\delta_A - \delta_B) - D_A(\omega_A - \omega_S) \tag{13}$$

$$\frac{1}{\omega_0}\dot{\delta}_A = \omega_A - \omega_S \tag{14}$$

IP turbine equations are:

$$2H_I\dot{\omega}_I = T_{IP} + K_{HI}(\delta_H - \delta_I) - K_{IA}(\delta_I - \delta_A) - D_I(\omega_I - \omega_S) \tag{15}$$

$$\frac{1}{\omega_0}\dot{\delta}_I = \omega_I - \omega_S \tag{16}$$

HP turbine equations are:

$$2H_H\dot{\omega}_H = T_{HP} - K_{HI}(\delta_H - \delta_I) - D_H(\omega_H - \omega_S) \tag{17}$$

$$\frac{1}{\omega_0}\dot{\delta}_H = \omega_H - \omega_S \tag{18}$$

In mechanical system equations, there are two types of damping, one is self damping and another is mutual damping. In this paper both type of damping is kept zero to show the effectiveness of STATCOM. Therefore in mechanical system equations only self damping terms are included and mutual damping terms are ignored for the simplicity.

Transmission Line Equations

$$V_{1D} - V_{2D} - V_{cD} = R_L(I_D + I_{sD}) - \omega X_{TL}(I_Q + I_{sQ}) + \frac{X_{TL}}{\omega_0}(\dot{I}_D + \dot{I}_{sD})$$

$$V_{1D} = V_{2D} + V_{cD} + R_L(I_D + I_{sD}) - \omega X_{TL}(I_Q + I_{sQ}) + \frac{X_{TL}}{\omega_0}(\dot{I}_D + \dot{I}_{sD}) \tag{19}$$

$$V_{1Q} - V_{2Q} - V_{cQ} = R_L(I_Q + I_{sQ}) + \omega X_{TL}(I_D + I_{sD}) + \frac{X_{TL}}{\omega_0}(\dot{I}_Q + \dot{I}_{sQ}) \tag{20}$$

$$V_{1Q} = V_{2Q} + V_{cQ} + R_L(I_Q + I_{sQ}) + \omega X_{TL}(I_D + I_{sD}) + \frac{X_{TL}}{\omega_0}(\dot{I}_Q + \dot{I}_{sQ})$$

$$\frac{\dot{V}_{cD}}{\omega_0} = \omega V_{cQ} + X_C(I_D + I_{sD}) \tag{21}$$

$$\frac{\dot{V}_{cQ}}{\omega_0} = -\omega V_{cD} + X_C(I_Q + I_{sQ}) \tag{22}$$

where, X_{TL} is the effective reactance of transmission line i.e. $X_{TL} = X_{Ts} + X_L$. X_L is sum of both 0.5 and 0.06. V_{cD} and V_{cQ} are the voltage across series capacitor. I_D and I_Q are D-axis and Q-axis component of current supplied by generator, similarly I_{sD} and I_{sQ} are the currents supplied by STATCOM.

STATCOM Circuit Equations

$$E_{sD} - V_{1D} = R_s I_{sD} - \omega X_s I_{sQ} + \frac{X_s \dot{I}_{sD}}{\omega_0} \tag{23}$$

$$E_{sQ} - V_{1Q} = R_s I_{sQ} + \omega X_s I_{sD} + \frac{X_s \dot{I}_{sQ}}{\omega_0} \tag{24}$$

$$E_{sD} = K_{cs} V_{dc} \cos\theta_s, \; E_{sQ} = K_{cs} V_{dc} \sin\theta_s$$

Linearization of E_{sD} and E_{sQ} is given here:

$$\Delta E_{sD} = K_{cs} \cdot \Delta V_{dc} \cdot \cos\theta_{s0} + K_{cs} V_{dc0}(-\sin\theta_{s0})\Delta\alpha$$
$$\Delta E_{sQ} = K_{cs} \cdot \Delta V_{dc} \cdot \sin\theta_{s0} + K_{cs} V_{dc0} \cos\theta_{s0}\Delta\alpha$$

STATCOM DC side power and AC side active power at node E_s must be equal. In steady state:

$$V_{dc} I_{dc} = \text{Real}\left[(E_{sD} + jE_{sQ})(I_{sD} + jI_{sQ})^*\right] = E_{sD} I_{sD} + E_{sQ} I_{sQ} \quad \text{or,}$$
$$V_{dc} I_{dc} = K_{cs} V_{dc} \cos\theta_s I_{sD} + K_{cs} V_{dc} \sin\theta_s I_{sQ}$$
$$\text{or,} \quad I_{dc} = K_{cs} \cos\theta_s I_{sD} + K_{cs} \sin\theta_s I_{sQ} \tag{25}$$

This is the expression of DC side current. In steady state numerical value of I_{dc} should be negative because R_p is consuming DC power, while I_{dc} shown in Fig. 1 is in opposite direction. Current through DC side capacitor

$$= C_s \dot{V}_{dc} = -\left(I_{dc} + \frac{V_{dc}}{R_p}\right) \quad \text{or,}$$
$$\dot{V}_{dc} = -\frac{1}{C_s}\left(\frac{V_{dc}}{R_p} + K_{cs} I_{sD} \cos\theta_s + K_{cs} I_{sQ} \sin\theta_s\right) \tag{26}$$

Linearization of Eq. 26 is as follows:

$$\Delta\dot{V}_{dc} = -\frac{1}{C_s}\left(\frac{\Delta V_{dc}}{R_p} + K_{cs}\Delta I_{sD} \cos\theta_{s0} - K_{cs}\Delta I_{sD0} \sin\theta_{s0}\Delta\theta_s + K_{cs}\Delta I_{sQ} \sin\theta_{s0} + K_{cs}\Delta I_{sQ0} \cos\theta_{s0}\Delta\theta_s\right).$$

PI Controller Equation

$$\dot{x} = V_{SS} + V_{ref} - |V_1| + K_d I_{sQ} \tag{27}$$

Output equation of PI controller is:

$$\alpha = K_i x + K_p \dot{x} \tag{28}$$

I_d and I_q are the output of generator system. I_d and I_q are converted to I_D and I_Q and then I_D and I_Q are input to transmission line system.

I_d and I_q are converted to I_D and I_Q with the following transformation

$$I_D = I_q \cos\delta + I_d \sin\delta, I_Q = I_q \sin\delta - I_d \cos\delta$$

Similarly $\dot{I}_d$ and $\dot{I}_q$ are needed to be changed in $\dot{I}_D$ and $\dot{I}_Q$. These quantities are also changed with above transformation. V_{1D} and V_{1Q} are the output of transmission line system and transformed to the e_d and e_q which are input to generator system. V_{1D} and V_{1Q} are transformed to e_d and e_q with the following transformation $e_d = V_{1D} \sin\delta - V_{1Q} \cos\delta, e_q = V_{1D} \cos\delta + V_{1Q} \sin\delta$

All the differential equations are developed in simulink and linearized in simulink as shown in Fig. 6. In Fig. 6 subsystem 'FBM' contains twenty differential equations, subsystem 'STATCOM' contains three differential equations. By keeping all the differential equations in a subsystem these can be linearized and subsequently eigenvalues, impulse response, root locus, observability etc. can be obtained.

Supplementary Controller Design

Various supplementary signals are tested in this study such as deviation in speed of generator rotor mode (ω–ω_s), deviation in reactive power, deviation in active power, deviation in generator current magnitude, deviation in STATCOM current magnitude and deviation in STATCOM DC side capacitor current ($C_s \dot{V}_{dc}$). It is found that deviation in speed of generator rotor mode ω–ω_s is most suitable supplementary signal. To find the best supplementary signal whole system is kept in the subsystem which input is V_{ref} and output is the candidate of supplementary signal. This system is without feedback supplementary controller and called 'sys1' hereafter. In each case, sys1 is linearized and pole-zero information is obtained. In any control system, poles (eigenvalues) are not dependent on choice of output, but zeros depends on the choice of output. Similarly, In this case poles of 'sys1' are same in each case, but zeros in each case are different. These zeros provide important information about the selection of supplementary signal. Root locus of a control system starts from poles of Open Loop Transfer Function (OLTF) and terminates at zeros of OLTF and it provides the information about the poles of closed loop transfer function. 'sys1' is OLTF as it is without feedback supplementary signal. The output which has less number of zeros in right hand side (RHS) of s-plane that is the best choice of supplementary signal (because if more number of zeros are in RHS, then root locus may be forced to shift in RHS). No. of zeros in RHS in each case are tabulated in Table 1.

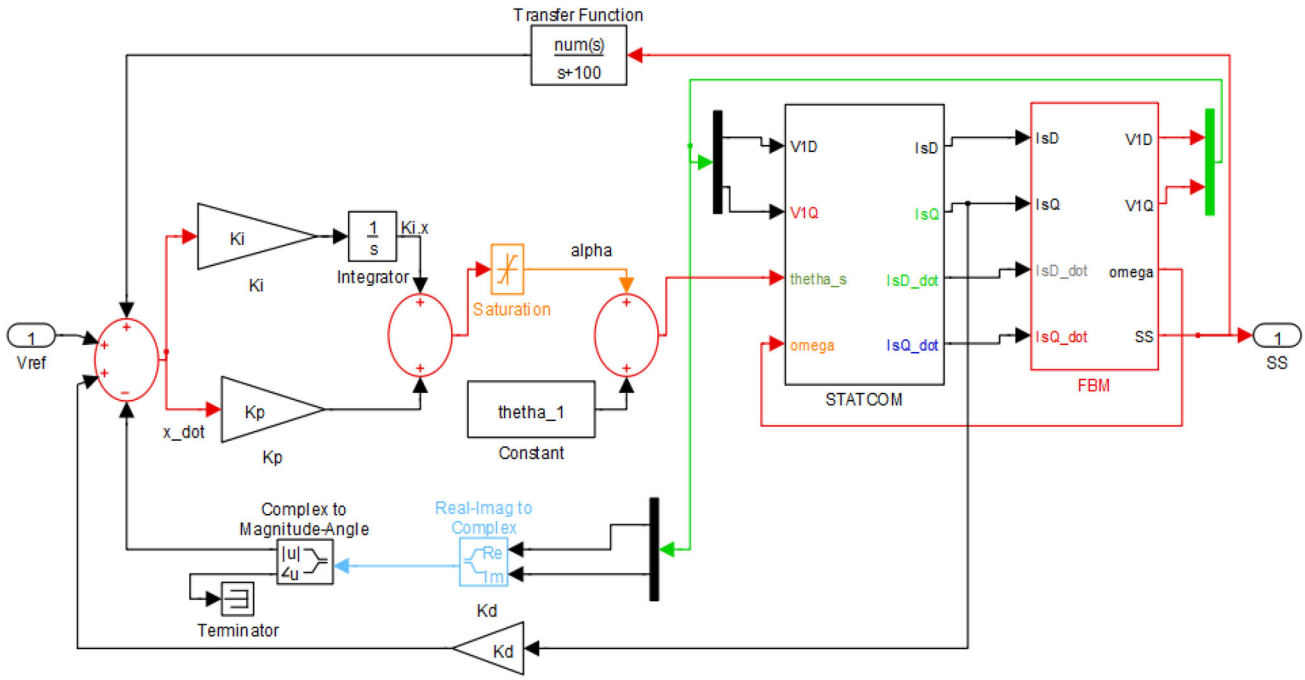


Fig. 6 Simulink model to solve all differential equations

From the results given in Table 1 it is found that deviation in speed of generator rotor mode is suitable signal. Apart from this eigenvalue analysis of each signal is carried out and it is found that only deviation in speed of generator rotor is able to damp all the modes. Optimum values of supplementary controller (i.e. K, z and p) and PI controller (i.e. K_p and K_i) is designed with the help of sisotool in MATLAB. Eigenvalues against designed supplementary controller are presented in Tables 2, 3, 4 and 5. Table 2 shows the eigenvalues when $X_c = 0.185$ pu. This critical value of X_c makes the Torsional mode 4 most unstable, because SSR frequency (subsynchronous frequency) is also in the range of 202. When supplementary signal ω–ω_s along with designed supplementary controller is incorporated then all eigenvalues are shifted to left half of s-plane. It can be seen that only STATCOM (with PI controller) cannot make the system stable. Similarly Table 3 shows the eigenvalues when $X_c = 0.284$ pu. This critical value of X_c makes the Torsional mode 3 most unstable, because SSR frequency is also in the range of 161. In this case also supplementary controller makes the system stable. Table 4 shows the eigenvalues at $X_c = 0.379$ pu. This critical value of X_c tends the Torsional mode 2 most unstable, because SSR frequency is also in the range of 126. In this case also supplementary controller makes the system stable. Table 5 shows the eigenvalues when $X_c = 0.473$ pu. This critical value of X_c makes the Torsional mode 1 most unstable, because SSR frequency is also in the range of 98. Here, in this case also supplementary controller makes the system completely stable. Time responses of deviation in speed of all the masses are shown in Figs. 7, 8, 9, 10, 11 and 12, Shaft torque between the masses are shown in Figs. 13, 14, 15, 16 and 17. Generator terminal voltage, active and reactive power supplied by generator is shown in Figs. 18, 19 and 20. Time response is shown against $X_c = 0.473$ pu. Disturbance is simulated after 1 s, which is twenty percent sudden increase in mechanical torque for 0.1 s. All the initial conditions are given in "Appendix" section.

Table 1 No. of Zeros in RHS of s-plane against different outputs

Output of 'sys1'	No. of zeros in RHS of s-plane
Deviation in speed of generator rotor mode	4
Deviation in reactive power	12
Deviation in active power	7
Deviation in generator current magnitude	7
Deviation in STATCOM current magnitude	8
Deviation in STATCOM DC side capacitor current	9

Conclusion

The modeling of a STATCOM with IEEE first benchmark model is presented in this paper. All the differential equations are developed in simulink, therefore linear analysis, root locus, pole—zero information etc. can be

Table 2 Eigenvalues at series compensation X_c = 0.185 pu, P_g = 0.9 pu, PF = 0.9 (lagging)

	Eigenvalues of IEEE first benchmark model only	Eigenvalues with STATCOM and PI controller	Eigenvalues with supplementary controller
Supersynchronous	− 4.5966 ± 551.17i	− 5.2224 ± 550.52i	− 5.5173 ± 550.21i
Torsional mode no. 5	− 1.204e−005 ± 298.18i	6.9031e−006 ± 298.18i	− 8.6287e−006 ± 298.18i
Torsional mode no. 4	1.9213 ± 202.43i	1.0048 ± 202.91i	− 4.824 ± 205.71i
Torsional mode no. 3	0.049347 ± 160.77i	0.050133 ± 160.78i	− 0.22416 ± 160.41i
Torsional mode no. 2	0.0057469 ± 127.05i	0.0075146 ± 127.05i	− 0.039108 ± 127i
Torsional mode no. 1	0.03539 ± 99.361i	0.066053 ± 99.448i	− 0.32946 ± 99.067i
Subsynchronous	− 5.6637 ± 202.85i	− 11.427 ± 201.72i	− 5.3645 ± 201.24i
Electromechanical mode	− 0.41516 ± 9.7111i	− 1.2538 ± 10.352i	− 1.2709 ± 10.276i
STATCOM currents		− 129.73 ± 2286.6i	− 130.3 ± 2286.3i
V_{dc}		− 224.12	− 217.95
Lead compensation			− 103.22
Others	− 32.493	− 32.844	− 33.103
	− 20.339	− 21.347	− 21.308
	− 3.3476	− 7.1655	− 7.1608
	− 0.26487	− 2.4311	− 2.433
		− 0.65913	− 0.65782

Table 3 Eigenvalues at series compensation X_c = 0.284 pu, P_g = 0.9 pu, PF = 0.9 (lagging)

	Eigen values of IEEE first benchmark model only	Eigen values with STATCOM and PI controller	Eigen values with supplementary controller
Supersynchronous	− 4.6347 ± 592.83i	− 5.9845 ± 592.03i	− 6.3268 ± 591.71i
Torsional mode no. 5	− 1.0878e−005 ± 298.18i	7.4748e−006 ± 298.18i	− 5.9055e−006 ± 298.18i
Torsional mode no. 4	− 0.060231 ± 202.64i	0.16974 ± 202.68i	− 0.40593 ± 204.12i
Torsional mode no. 3	1.8221 ± 160.36i	0.69405 ± 160.65i	− 1.7819 ± 160.68i
Torsional mode no. 2	0.021412 ± 127.1i	0.019667 ± 127.11i	− 0.050933 ± 126.94i
Torsional mode no. 1	0.092919 ± 99.673i	0.037375 ± 99.759i	− 0.24589 ± 98.9i
Subsynchronous	− 5.127 ± 161i	− 14.56 ± 158.29i	− 11.576 ± 160.3i
Electromechanical mode	− 0.51673 ± 10.756i	− 1.3479 ± 11.035i	− 1.3544 ± 10.953i
STATCOM currents		− 129.72 ± 2286.7i	− 130.29 ± 2286.5i
V_{dc}		− 216	− 211.57
Lead Compensation			− 101.52
Others	− 32.715	− 33.017	− 33.303
	− 20.382	− 21.32	− 21.279
	− 3.5521	− 7.0403	− 7.0359
	− 0.31799	− 2.5828	− 2.5848
		− 0.66126	− 0.65997

easily obtained. In this paper, it is found that supplementary signal, deviation in speed of generator rotor mode, is a very suitable supplementary signal for STATCOM which can make the system stable against all four critical values of X_c. In the present study all the self damping and mutual damping are kept zero to show the effectiveness of proposed controller. The eigenvalues analysis and the nonlinear time domain simulation results show the effectiveness and the robustness of the proposed controller. Transient responses of each shaft torque and speed deviation of all the masses are presented. The proposed controller is easy to implement in the power systems.

Appendix

Initial conditions of the system is given here. All the quantities are in pu unless otherwise specified.

Table 4 Eigenvalues at series compensation X_c = 0.379 pu, P_g = 0.9 pu, PF = 0.9 (lagging)

	Eigen values of IEEE first benchmark model only	Eigen values with STATCOM and PI controller	Eigen values with supplementary controller
Supersynchronous	− 4.6613 ± 626.34i	− 6.6704 ± 625.51i	− 7.0387 ± 625.19i
Torsional mode no. 5	− 1.0164e−005 ± 298.18i	8.1528e−006 ± 298.18i	− 3.698e−006 ± 298.18i
Torsional mode no. 4	− 0.047217 ± 202.76i	0.1209 ± 202.76i	− 0.30037 ± 203.9i
Torsional mode no. 3	− 0.040358 ± 160.29i	0.21139 ± 160.4i	− 0.4613 ± 161.28i
Torsional mode no. 2	0.86533 ± 126.81i	0.20415 ± 127.04i	− 0.32336 ± 126.99i
Torsional mode no. 1	0.33001 ± 100.56i	0.43002 ± 100.61i	− 0.35991 ± 98.23i
Subsynchronous	− 3.8539 ± 126.76i	− 17.947 ± 121.42i	− 16.229 ± 125.43i
Electromechanical mode	− 0.65265 ± 11.881i	− 1.4999 ± 11.821i	− 1.4906 ± 11.727i
STATCOM currents		− 128.39 ± 2288.7i	− 128.94 ± 2288.5i
V_{dc}		− 206.19	− 203.68
Lead Compensation			− 99.005
Others	− 33.005	− 33.103	− 33.427
	− 20.44	− 21.313	− 21.268
	− 3.8364	− 6.9611	− 6.9567
	− 0.36962	− 2.7121	− 2.7143
		− 0.66245	− 0.66116

Table 5 Eigenvalues at series compensation X_c = 0.473 pu, P_g = 0.9 pu, PF = 0.9 (lagging)

	Eigenvalues of IEEE first benchmark model only	Eigenvalues with STATCOM and PI controller	Eigenvalues with supplementary controller
Supersynchronous	− 4.6828 ± 655.56i	− 7.2881 ± 654.78i	− 7.6675 ± 654.48i
Torsional mode no. 5	− 9.4572e−006 ± 298.18i	8.7833e−006 ± 298.18i	− 6.1047e−007 ± 298.18i
Torsional mode no. 4	− 0.039127 ± 202.81i	0.10989 ± 202.8i	− 0.26315 ± 203.83i
Torsional mode no. 3	− 0.031749 ± 160.42i	0.1382 ± 160.46i	− 0.31163 ± 161.15i
Torsional mode no. 2	− 0.010672 ± 126.9i	0.083774 ± 126.97i	− 0.16357 ± 127.15i
Torsional mode no. 1	5.7345 ± 98.456i	2.0164 ± 99.482i	− 2.6406 ± 98.883i
Subsynchronous	− 7.6935 ± 99.15i	− 22.736 ± 89.13i	− 18.592 ± 92.803i
Electromechanical mode	− 0.86108 ± 13.24i	− 1.7657 ± 12.816i	− 1.7229 ± 12.698i
STATCOM currents		− 125.94 ± 2291.8i	− 126.48 ± 2291.6i
V_{dc}		− 194.46	− 194.48
Lead Compensation			− 94.73
Others	− 33.415	− 33.038	− 33.416
	− 20.522	− 21.315	− 21.266
	− 4.2366	− 6.88	− 6.8756
	− 0.42982	− 2.8316	− 2.8338
		− 0.66335	− 0.66204

$V_1 = 1.009\angle 11.818^\circ$, $V_2 = 0.963\angle 0$,
$E_s = 0.975\angle 11.918^\circ$, $\alpha = 0.1^\circ$.

Current supplied by generator = 0.962 − j0.245, Power supplied by generator = 0.899 + j0.441, Current supplied by STATCOM = − 0.049 + j0.219, Power supplied by STATCOM = − 0.004 − j0.226 (i.e. STATCOM is consuming active power and reactive power, therefore it is in inductive mode). STATCOM circuit data: R_s = 0.01, X_s = 0.15, R_p = 125, C_s = 0.00059, K_{cs} = 1.559

Mechanical system initial conditions:

$\delta_E = \delta = 52.77^\circ$, $\delta_B = 53.49^\circ$, $\delta_A = 54.27^\circ$, $\delta_I = 55.10^\circ$, $\delta_H = 55.90^\circ$

PI controller data: K_p = − 1, K_i = − 3.75 s^{-1}. K_d = 0.01. Data of transfer function in feedback

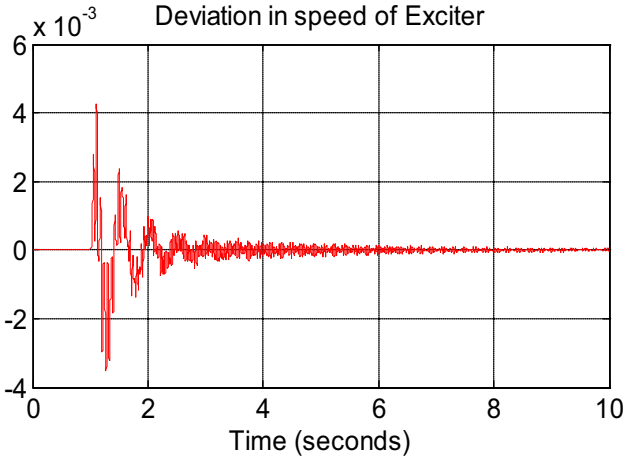


Fig. 7 Time response of deviation in speed of exciter

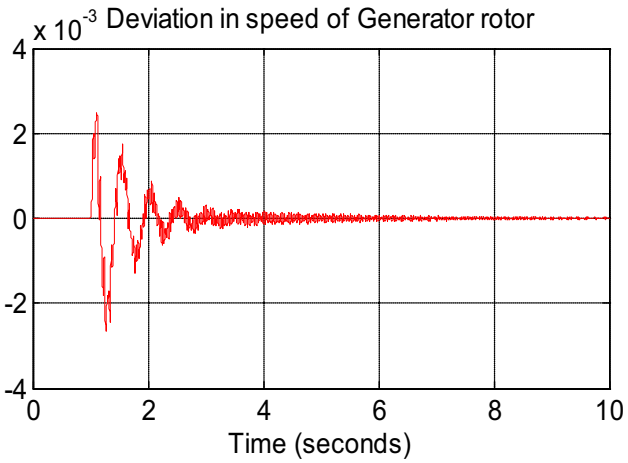


Fig. 8 Time response of deviation in speed of generator rotor

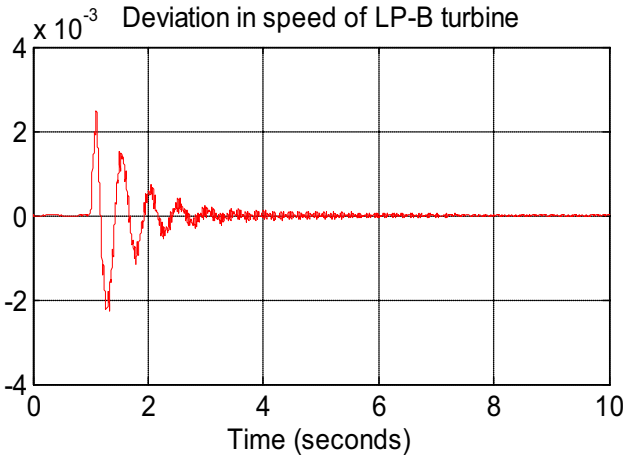


Fig. 9 Time response of deviation in speed of LP-B turbine

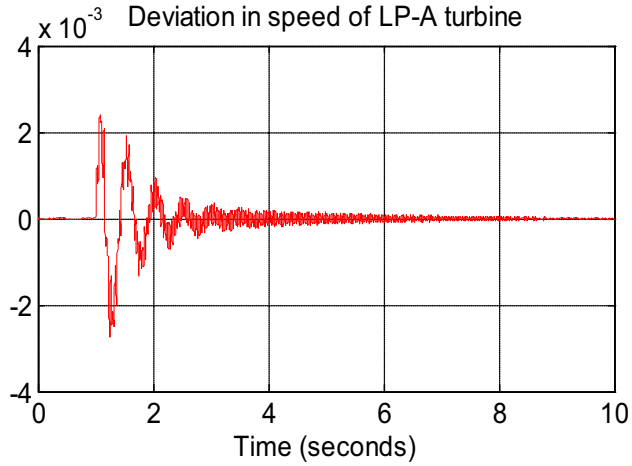


Fig. 10 Time response of deviation in speed of LP-A turbine

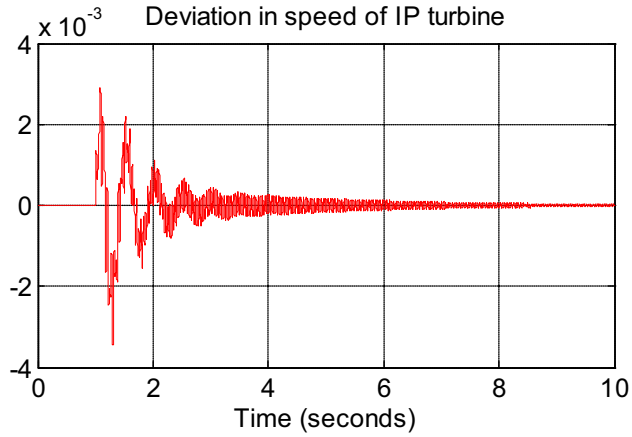


Fig. 11 Time response of deviation in speed of IP turbine

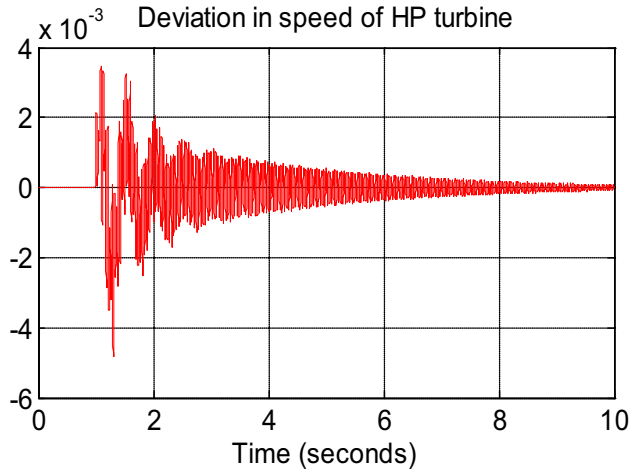


Fig. 12 Time response of deviation in speed of HP turbine

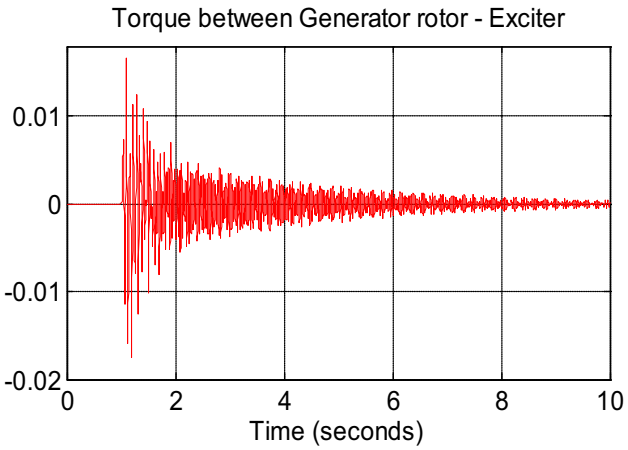


Fig. 13 Time response of Torque between generator rotor–exciter

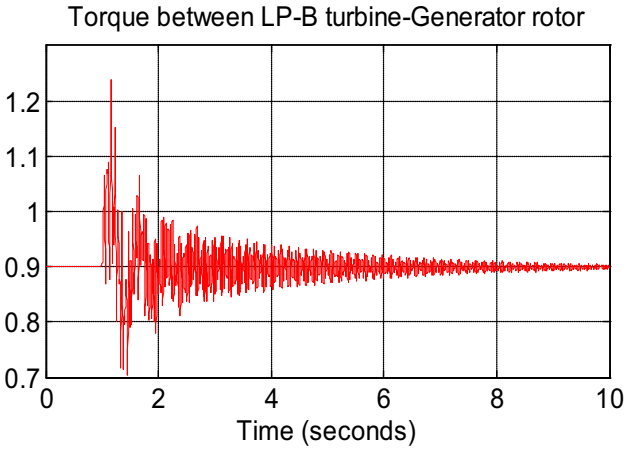


Fig. 14 Time response of Torque between LP-B turbine Generator rotor

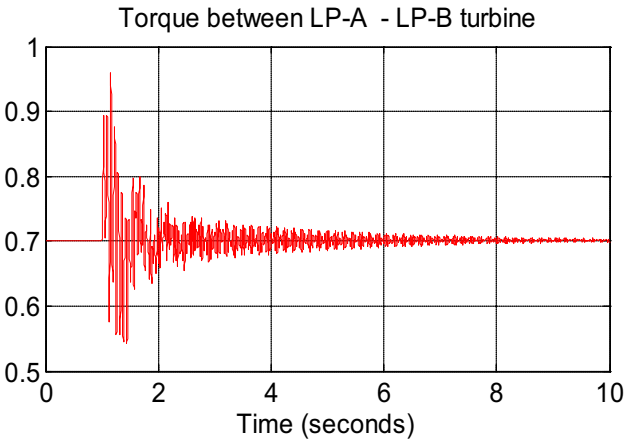


Fig. 15 Time response of Torque between LP-A turbine-LP-B turbine

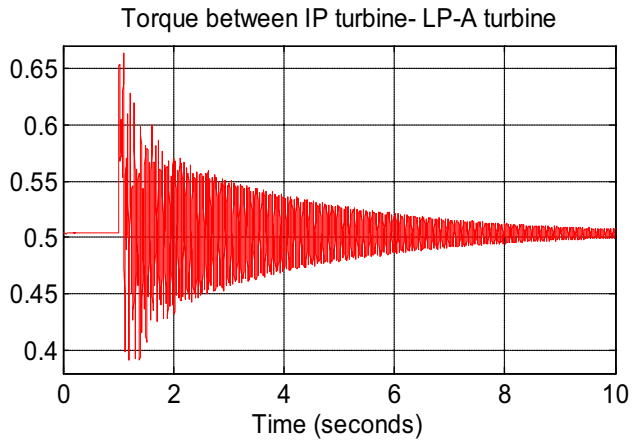


Fig. 16 Time response of Torque between IP turbine-LP-A turbine

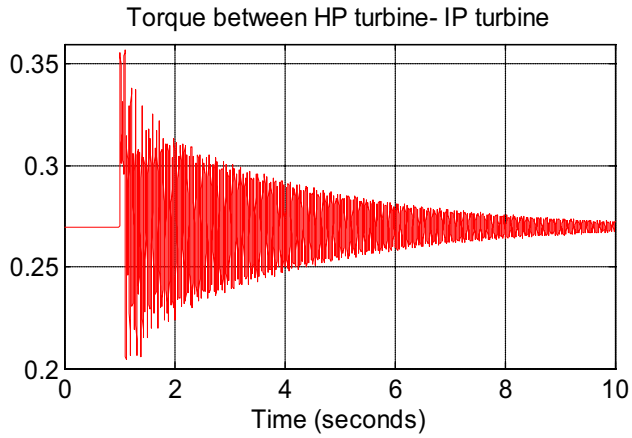


Fig. 17 Time response of Torque between HP turbine–IP turbine

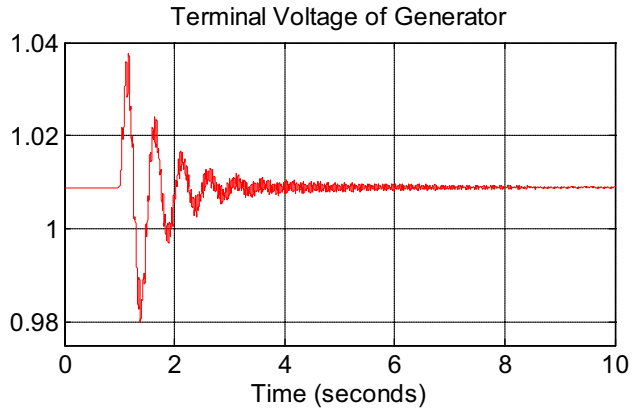


Fig. 18 Time response of terminal voltage of generator

supplementary signal path: K = 15, z = 6 s^{-1}, p = 100 s^{-1}. In Fig. 2 unit of 'x' is seconds and $\dot{x}$ is pu. H_E, H_G, H_B, H_A, H_I, H_H are inertia constants of different masses. ω_E, ω, ω_B, ω_A, ω_I, ω_H are the speed of different masses. ω_S is synchronous speed. In steady state

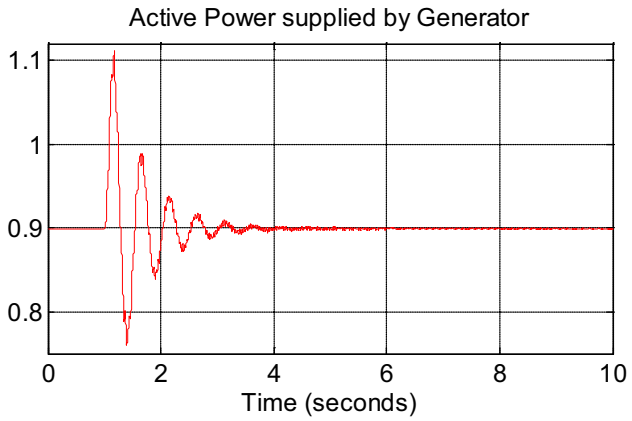


Fig. 19 Time response of active power supplied by generator

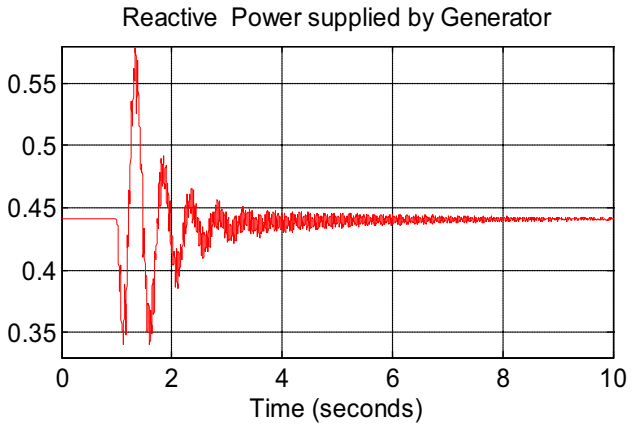


Fig. 20 Time response of reactive power supplied by generator

$\omega_E = \omega = \omega_B = \omega_A = \omega_I = \omega_H = \omega_S = 1$ pu. In transient period ω_S remains 1 pu. $\omega_0 = 376.99$ rad/s. Its value is also fixed in transient period. Generator circuit data, mechanical system data are given in [18].

References

1. N. Kumar, M.P. Dave, Application of static VAR system supplementary controllers to improve the transient performance of series compensated long transmission lines. Electr. Power Syst. Res. **34**(2), 75–83 (1995)
2. M.R. Banaei, H. Toloue et al., Damping of power system oscillations using imperialist competition algorithm in power system equipped by HVDC. Ain Shams Eng. J. **6**(1), 75–84 (2015)
3. A. Safari, A. Ahmadian, M.A.A. Golkar, Controller design of STATCOM for power system stability improvement using honey bee mating optimization. J. Appl. Res. Technol. **11**, 144–155 (2013)
4. S.T. Nagarajan, Investigations on AI based control of FACTS devices for damping SSR in power system, Thesis, Delhi College of Engineering, University of Delhi, Delhi, India (2013)
5. K.R. Padiyar, *FACTS Controllers in Power Transmission and Distribution* (New Age International Publishers, New Delhi, 2007), pp. 190–195
6. K.V. Patil, Dynamic compensation of electrical power systems using a new BVSI STATCOM, Thesis, The University of Western Ontario, London, Ontario (1999)
7. D. Soto, R. Pena, Nonlinear control strategies for cascaded multilevel STATCOMs. IEEE Trans. Power Del. **19**(4), 1919–1927 (2004)
8. R.M. Mathur, R.K. Varma, *Thyristor Based Facts Controllers for Electrical Transmission Systems* (Wiley, London, 2002), pp. 420–427
9. K.V. Patil, J.S.J. Jiang, R.M. Mathur, Application of STATCOM for damping torsional oscillations in series compensated AC systems. IEEE Trans. Energy Conver. **13**(3), 237–243 (1998)
10. P. Rao, M.L. Crow, Z. Yang, STATCOM control for power system voltage control applications. IEEE Trans. Power Deliv. **15**(4), 1311–1317 (2000)
11. A. Jain, K. Joshi et al., Voltage regulation with STATCOMs: modeling, control and results. IEEE Trans. Power Del. **21**(2), 726–735 (2006)
12. D. Rai, Damping subsynchronous resonanceusing static synchronous series compensators and static synchronous compensators, Thesis, University of Saskatchewan, Saskatoon, Canada (2008)
13. D. Rai, S.O. Faried et al., Impact of imbalanced phase operation of SSSC on damping subsynchronous resonance, in *Power and Energy Society General Meeting* (2011)
14. J.K. Moharana, M. Sengupta, A. Sengupta, Design, analysis and implementation of a small signal control strategy on a 10 KVA STATCOM prototype connected to inductive load. J. Inst. Eng. India Ser. B **94**(1), 61–69 (2013)
15. M.T. Bina, D.C. Hamill, Average circuit model for angle-controlled STATCOM. IEE Proc. Electr. Power Appl. **152**(3), 653–659 (2005)
16. L. Xing, A comparison of pole assignment and LQR design methods for multivariable control for STATCOM, Thesis, The Florida State University (2003)
17. K.S. Jeong, T.Y. Jyung, Y.S. Baek, Modeling and dynamic analysis of STATCOM for short-term voltage stability improvement using EMTP/RV. J. Int. Counc. Electr. Eng. **1**(2), 129–134 (2011)
18. IEEE Subsynchronous Resonance Task Force, First benchmark model for computer simulation of subsynchronous resonance. IEEE Trans. Power Appar. Syst. **PAS-96**(5), 1565–1572 (1977)
19. S. Gomes, P.E.M. Quintao, N. Martins, Subsynchronous resonance results obtained with a comprehensive small signal stability program, in IX Symposium of Specialists in Electric Operational and Expansion Planning (2004), p. 1–9
20. G. Tang, Damping subsynchronous resonance oscillations using a VSC HVDC back-to-back system, Thesis, University of Saskatchewan, Saskatoon, Saskatchewan (2006)
21. R. Satija, Damping subsynchronous resonance in power system using Static VAR system, Thesis, University of Delhi, Delhi (2007)

Digital Video Watermarking: Issues and Challenges

Md Shahid and Pradeep Kumar

Abstract: A pilfered duplicate of a digital video would be easily disseminated to the global audience because of the rapid high-speed internet. Due to impeccably replicable nature of digital video, numerous unlawful duplicates of the original video can be made. A video can undergo several intentional attacks like frame dropping, averaging, cropping and median filtering and unintentional attacks like the addition of noise and compression which can compromise copyright information, thereby denying the authentication. Hence techniques are needed to secure copyrights of the proprietor and counteract illegal copying. One of the techniques is Video Watermarking strategy for concealing some sort of information into digital video sequences that are orders of successive still frames. In this paper, we study properties of video watermarking, the arrangement of computerized video watermarking systems, watermark attacks, its applications, issues and challenges for video watermarking. At last, we propose some future research directions.

***Index Terms*—** Watermarking techniques, Attacks, LSB, DWT, DCT, DFT, FFT, SVD.

I. INTRODUCTION

A large amount of digital multimedia data is available today, which can be perfectly copied and rapidly disseminated at large scale. This consequently has raised concerns from the content owners, when they realized that traditional protection mechanisms, such as encryption, were no longer sufficient. Sooner or later, digital content has got to be decrypted and to be presented to users. At this very moment, the protection offered by encryption no longer exists. As a result, digital watermarking has been studied as a complementary technology. After Image watermarking scientist invented video watermarking [1, 2]. Watermarking is a process of embedding copyright information into the cover-data, also called host- data in order to protect the intellectual right and the originality of cover data. The host-data can be text, audio, video, etc. It is just an extension of the watermarking concept [3]. With the same and optimized length of the video, video watermarking allows embedding more data without degrading
the quality of the host-video much. A digital video is made of a sequence of still images/frames. Video watermarking hides the data-information about the authenticity or integrity of host- video in the frames of host-video. One can either select some frames of the host video or all to embed a watermark into video. This process of embedding watermark into video is called video watermarking [1]. Both of the data hiding techniques- Steganography as well as watermarking use similar techniques to hide information into host-data. And

both of them are used to hide information but their goals are different. In Digital Watermarking the content owner is concerned about the quality of host-data as well as integrity of watermark. In Steganography, the sender and the recipient of the data are more concerned about integrity of hidden information only, which is encoded in host-data and it is hidden only unlike watermark which can even be visible to the public.

2. General Properties of Video Watermarking

Some of the general properties of video watermarking are given below, which play a very significant role in video watermarking:

2.1 Imperceptibility:

The watermark should be imperceptible in such a way that third person should not be able to find out whether a video is watermarked or not [4].

2.2 Robustness:

Whenever a video is shared there are some distortions. The watermark should be robust against malicious attacks in such a way that even if the cover video changes, copyright data, which is embedded into cover video, should not get affected [5, 6].

2.3 Time complexity and Computational cost:

Cost of embedding a watermark into host- data and to extract the embedded watermark from the cover data should be less [4].

2.4 Capacity and Payload:

The amount of data to be embedded in cover data or host data is called capacity. Payload is the number of watermark bits in the cover data. Number of watermark bits in the host-data varies from one application to other [5, 7].

2.5 Security:

The watermark and original data should not be accessible to the unauthorized user. Watermark and original data should not be affected by any form of attacks [4, 5].

3. CLASSIFICATION OF VIDEO WATERMARKING TECHNIQUE

In the Fig 2, classification of video watermarking is shown which has been done depending upon perception and domain.

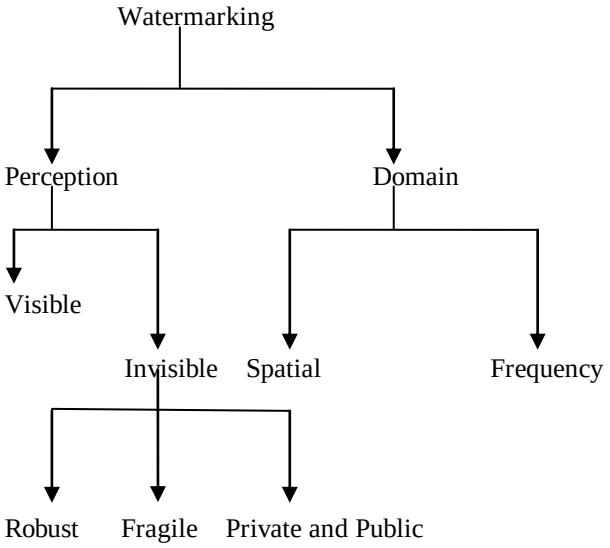


Fig 2: Classification of Video Watermarking Techniques.

3.1 Spatial Domain Method (SDM):

In the spatial domain method, a watermark is embedded by modifying the pixel value; which is made of digital information of the host video [4, 9]. This method is faster as compared to Frequency domain in simplicity and less computational complexity [5]. This method is less robust to geometric attacks and less resistant to noise, compression and low pass filtering [6].

3.1.1 Least Significant Bit (LSB):

It is simplest watermarking technique; the watermark is directly embedded by changing some of the lower-oder bits of each pixel. The amount of watermark data in LSB technique is very less and restricted [7]. The least significant bit (LSB) technique is used for simple operation to embed information in a cover image or frame. The LSB technique is that inside of a cover, image pixels are changed by bits of secret message.

3.1.2 Correlation -Based Technique:

It is one of the spatial domain techniques of watermarking, in this approach of digital watermarking; the watermark is embedded by incorporating pseudorandom noise pattern to luminance value of video pixels. The pseudorandom noise (MN) pattern i.e. A(y, z) is added to host- image B(y, z) the outcome is watermarked image BA(y, z). Here (y, z) represent position of watermark. $BA(y, z) = B(y, z) + k^* A(y, z)$. k is the gain factor. [6].

3.2 Frequency Domain Method (FDM):

In the frequency domain method, the watermark is embedded into host-data after converting digital information into frequency using transformation tools [4]. This method of watermarking is more robust, efficient and secure as compared to spatial domain [6].

3.2.1 Discrete Fourier Transform (DFT):

In this approach, a watermark is embedded only in the first frame of a group of pictures. The full DFT technique is applied to identify frame to be watermarked and magnitude of the coefficient is calculated [6].

3.3.1 Singular Value Decomposition (SVD):

It is used to get diagonalized matrices. In this technique, singular value of original image is used to embed the watermark. The length and coefficient of watermark effect watermarking technique [6]. SVD is one of the mathematical tools to analyze matrices. A square matrices A of M x N size is decomposed into three matrices X,Y,Z such that A=XYZT. Where, ZT transpose of Z, X, Y are orthogonal matrix and z is square diagonal [12].

3.3.2 Discrete Cosine Transform (DCT):

This works as a tool to convert spatial domain into frequency domain. In DCT, data is represented in terms of frequency [7].
Embedding of watermark is performed in first k highest magnitude of DCT coefficient of image/frame. The result obtained after applying DCT coefficient will be an image of sum of varying magnitude and frequency. Horizontal frequencies vary from left to right and vertical frequencies top to bottom. Watermark in mid band gives better robustness to attacks and imperceptibility [4, 6].

3.3.3 Discrete Wavelet Transform (DWT):

This technique is based on wavelets concept with varying frequency with respect to duration. Multi-resolution decomposition of images and frames is possible in DWT. The wavelet is divided into 4 sub bands LL, LH, HL, and HH. The first letter in sub band refers to frequency applied for rows and second letter refers to filter applied to columns [13].

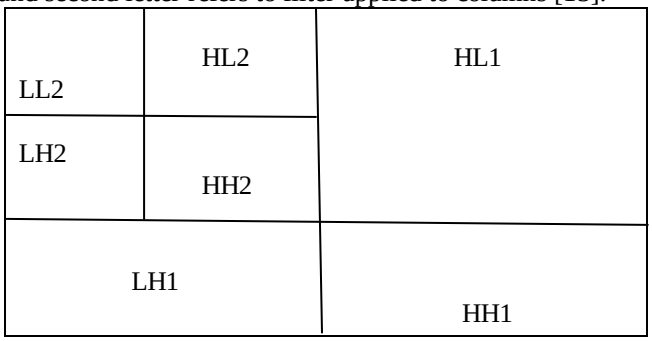


Fig 3: 2-levels DWT Scheme

All Rights Reserved © 2018 IJARCET

We present in a tabular form a comparative analysis of different standard watermarking techniques in terms of their features. Following are the expressions used in the table :

RO: Robustness, IM: Imperceptibility

SE: Security

PL: payload , CC: Computational Cost

TC: Time Complexity, RE: Reliability.

Features	LSB	DFT	DCT	DWT	SVD
RO	Less robust against geometric distortion	High robust against geometric distortion	High robust against filter	High robust against geometric distortion	High robust against geometric distortion
IM	Less compared to others.	High	High	Better watermark length and visual quality	Better length of co-efficient effect visual quality
SE	Less secure, usually depends on choice key	Highly secure	Better, semi private watermarking	Better, semi private watermarking	High private watermarking
P	Less, limited data can be added	Average	High	High	High
CC	Less	Reliable cost	Reliable cost	Very high	High
TC	Less	High	High	Very high	high
RE	Better for multiple watermark	High	high	Very high	Very high

4. VIDEO WATERMARKING APPLICATIONS

Following are the some applications of digital video watermarking:

- Source Tracking
- Video Authentication
- Copy-right protection
- Broadcast Monitoring of Video Sequences

5. WATERMARK ATTACKS

The existing class of attacks contains many attacks: e.g. simple attacks, geometric attacks, cryptographic attacks, protocol attacks, etc. Here, we describe some of these attacks types.

5.1 Active attacks:

Attackers can manipulate data and make it undetectable. But in the active attack of digital watermarking, the attacker tries deliberately to eliminate the watermark or simply make it undetectable. This type of attack is grave for many applications where the purpose of the watermark is of no use when it can't be detected.

5.2 Passive attacks:

In passive attacks, attacker does not try to remove the watermark but simply attempts to determine if a given mark is present or not. Protection against these kind of attacks is of the utmost importance in covert communications where the simple knowledge of the presence of watermark is often more than one wants to grant.

5.3 Geometric attacks:

These kinds of attacks just distort the watermark detector synchronization with the embedded data; it means these attacks do not remove the embedded watermark itself. [8, 21, 18].

5.4 Collusion attacks:

In these kinds of attacks, the attacker tries to remove the watermark as for the active attacks but the method is quite different. In order to eliminate the watermark, the attacker uses many copies of the same data, containing each different watermark, each signed with a key to construct a new copy without any watermark. These kinds of attacks are not so easy.

5.5 Forgery attacks:

In these kinds of attacks, the hacker's goal is to embed a new watermark rather than removing one. By doing so, it allows one to modify the protected data and then, re-implants a new

given key to replace the destructed one, thus making the corrupted image seems legit.

5.6 Simple Attacks:

In these kinds of attacks, the attacker attempts to impair the embedded watermark by manipulating the watermark and host data without any attempt to identify and isolate the embedded watermark.

6. RELATED WORK

Sanjai Kumar and Ambar Dutta reviewed in the paper, performance analysis of spatial domain digital watermarking techniques. where they implemented two spatial domain algorithms for digital watermarking and compared them with respect to different performance metrics in the presence and absence of different types of noise. From the experimental results they found that the algorithm using the concept of maximum entropy block in the cover image (Algorithm – 2) performed better as compared to Algorithm – 1. It was also found from the results that Algorithm – 1 is more robust with respect to Gaussian noise compared to Algorithm – 2. Both the algorithms are robust with respect to salt-and-pepper noise [20].

Namita Tiwari and Sharmila reviewed a paper , Digital Watermarking Applications, Parameter Measures and techniques[24]. They found that for security purpose transformation techniques are more suitable and Spatial domain techniques are easy to implement.

Revathi.B, Vijaya Geetha.R, Vengadapathiraj.M, Maheswaran.U and Karunakaran proposed a paper , digital video watermar king algorithm for content validation using singular value decomposition [25]. They found that the contrast in Table 1 method is good when associated to the existing watermarks. As an anticipated work, can go for inlaying diverse watermarks on the different frames of the content.

Naved Alam proposed a scheme in the paper, A Robust Video Watermarking Technique using DWT, DCT, and FFT[26]. It shows from the results that watermarked video sequences are very much susceptible to pirate attacks.

Chitrasen and Tanuja Kashyap proposed a technique in the paper [27], Robust Video Watermarking using Discrete Cosine Transform and Third level Discrete Wavelet Transform. They found that When the embedding strength is stronger, the robustness is greater but visual experience is worse. From the result, it shows that the proposed algorithm has a strong ability to resist different watermark attacks.

M. Sundararajan and G.Yamuna proposed a technique in the paper[28], dwt based scheme for video watermarking, from the results, it shows that the proposed scheme is efficient by means of imperceptibility and robustness.

Paramjit Kaur and Dr. Vijay Laxmi proposed a scheme in the paper[29], An Upgraded Approach for Robust Video Watermarking Technique Using Stephens Algorithm. From the results, it shows that embedded video is excellent with high PSNR, better efficiency and has very low visual artifacts.

Bhargavi Latha S, Venkata Reddy Dasari and Damodaram Avula proposed a scheme in the paper [30], Robust Video Watermarking Using Secret Sharing, SVD, DWT and Chaotic Firefly Algorithm. From the experimental results, it's found that the proposed technique of video watermarking is better in terms of performance when compared with state of art methods.

R.Lancini, F.Mapelli and S.Tubaro proposed a video watermarking technique in the paper [32], a robust video watermarking technique in the spatial domain. From the results, it shows that the proposed technique allows us to hide invisible data in a retrievable way.

Amir Houmansadr and Shahrokh Ghaemmaghami proposed a video watermarking technique in the paper[33], A Novel Video Watermarking Method Using Visual Cryptography. Frpm the reults, it shows that method hasa high resilience against geometric distortions and other non-hostile video processing schemes.

Chien-Chuan Ko and Bo-Zhi Yang proposed a technique of video watermaking in the paper[34],An Integrated Technique for Video Watermarking. From the results, it shows that this method can resist attacks such as JPEG compression, sharpening, median filter, blur, drop noise, brightness adjustment,cropping, Gaussian noise, and MPEG compression etc.

7. ISSUES AND CHALLENGES FOR VIDEO WATERMARKING

From the study of digital watermarking techniques, we have found that when a watermarked video is shared there is always a chance of being attacked. Therefore, during the design of a watermark algorithm, these are following some issues which need to be addressed properly:

- Capacity and Payload
- Robustness
- Transparency
- Security

8. CONCLUSION

This paper is an attempt to review digital video watermarking techniques, their properties, their applications, attacks and classifications. Reviewed literature showed that there are many types of digital video

All Rights Reserved © 2018 IJARCET

watermarking techniques that play crucial role in data protection in any form like images, audios, text and/or videos. Researchers used different types of equations in their literature to hide the information in the cover video, which is sequence of still frames. To further improve the video watermarking techniques and make the system more robust and secure from various attacks integration of various video watermarking techniques are required.

References:

[1] Yafeng Zhou and Wing W.Y. Ng Conference paper on influence between digital watermarking and steganography | IEEE, ISSN: 2158-5709, 16 December 2013.
[2] Jashandeep Kaur Kang, Rakesh Kumar, Kamaljeet Kainth, ―Review paper on video watermarking‖ International journal of advanced research in computer Science and software engineering, vol.6, issue 6, June 2016.
[3] Ankitha.A.Nayak et al. Int. Journal of Engineering Research and Applications .ISSN : 2248-9622, Vol. 4,Issue 12(Part 6), December 2014, pp.39-44
[4] Prabhishek Singh, R S Chadha "A Survey of Digital Watermarking Techniques, Applications and Attacks" International Journal of Engineering and Innovative Technology (IJEIT), March 2013.
[5] Sourav Bhattacharya, T. Chattopadhyay, Arpan Pal, ―A Survey onDifferent Video Watermarking Techniques and Comparative Analysis with Reference to H.264/AVC‖, IEEE 2006
[6] Gopika V Mane, G. G. Chiddarwar, ―Review Paper on Video Watermarking Techniques‖, International
Journal of Scientific and Research Publications,Volume 3, Issue 4, April 2013 1 ISSN 2250-3153
[7] Dr.V.Seenivasagam, S. Subbulakshmi, S. Radhamani,―A survey on video watermarking and its applications‖ International journal of advanced research in computer science and software engineering, vol 4,issue 3, march 2014
[8] Wiem Trabelsi, Mohamed Heny Selmi, ―Multi-signature robust video watermarking‖ 1st International
Conference on Advanced Technologies for Signal and Image Processing , Sousse, Tunisia, March 2014.
[9]Mustafa Osman Ali and Rameshwar Rao, "Digital Image Watermarking Basics, and Hardware Implementation", International Journal of Modeling and Optimization, Vol. 2, No. 1, February 2012.
[10] Gui Feng, Kai Huang "H.264 Video Standard Based Zero Watermarking Technology", project supported by the Natural Science Foundation of Fujian Province of China.
[11] Mrs. Anita Jadhav, Mrs. Megha Kolhekar, "Digital Watermarking in Video for Copy Right Protection",International Conference on Electronic Systems,Signal Processing and Computing Technologies,2014.
[12] Y.J.Song, T.N.Tan "Comparison of Four Different Digital Watermarking Techniques", Proceedings of
ICSP2000.
[13] Hitesh Panchal,Kunal Acharya, Pradip Panchal,Naimish Thakar "Digital Watermarking on Extracted Key Frames from Uncompressed Color Video using 4- Level DWT" , 3rd International Conference on Computer Engineering and Technology,2010.
[14] Swati Patel ,Anilkumar Katharotiya, Mahesh Goyani-A survey on digital video watermarking‖ Int. J.Comp. Tech. Appl., Vol 2 (6), 3015-3018
[15] T. L. Gilbert, Formulation, Foundations and Applications of the PhenomenologicalTheory Ferromagnetism, Ph.D. dissertation, Illinois Inst.Tech., Chicago, IL, 1956, unpublished.
[16] Mahima Jacob, Saurabh Mitra,‖ Video Watermarking Techniques‖, IJRTE, Vol 4, Pp 1- 4,2015
[17] Lalit Kumar Saini, Vishal Shrivastava,‖ A Survey of Digital Watermarking Techniques and itsApplications‖, IJCST, Vol 2, Pp 70-73,2014
[18 Rakesh Ahuja, S. S. Bedi,‖ All Aspects of Digital Video Watermarking Under an Umbrella‖, Ijigsp, Vol 12, Pp 54-73, 2015
[19] Amit Kumar Singh, Nomit Sharma, Mayank Dev and Ananad Mohan| An International Conference Paper, A novel technique for digital image watermarking in spatial domain, IEEE, 2012.
[20] Sanjai Kumar and Ambar Dutta| An International Conference paper, Performance analysis of spatial domain digital watermarking techniques, ISBN: 978-1-5090-2552-7,IEEE, 2016.
[21] Xing Chang, Weilin Wang, Jianyu Zhao, Li Zhang,"A Survey of Digital Video Watermarking",2011 Seventh International Conference on Natural Computation,61-65.
[22] Mehdi Fallahpour, Shervin Shirmohammadi, Mehdi Semsarzadeh,and Jiying Zhao, ―Tampering Detection in Compressed Digital Video Using Watermarking,‖ IEEE Transactions on Insrumentation and Measurement, vol. 63, no. 5, May 2014.
[23] Dolly Shukhla and Manish Sharma - Robust Scene-Based Digital Video Watermarking Scheme Using Level-3 DWT: Approach, Evaluation, and Experimentation | Spinger , January 2018, Volume 61, Issue 1, pp 1–12.
[24] Namita Tiwari and Sharmila - Digital Watermarking Applications, Parameter Measures and Techniques, IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.3,March 2017.
[25]Revathi.B, Vijaya Geetha.R, Vengadapathiraj.M, Maheswaran.U and Karunakaran- digital video watermar king algorithm for content validation using singular value decomposition, International Journal of Science, Engineering and Technology Research (IJSETR)Volume 5, Issue 4, April 2016.
[26]Naved Alam - A Robust Video Watermarking Technique using DWT, DCT, and FFT, Volume 6, Issue 6,June2016,International Journal of Advanced Research in Computer Science and Software Engineering.
[27] Chitrasen and Tanuja Kashyap - Robust Video Watermarking using Discrete Cosine Transform and Third level Discrete Wavelet Transform, Journal of Engineering Research and Application Vol. 7, Issue 10, (Part -1) October 2017, pp.87-92.
[28] M. Sundararajan and G.Yamuna- dwt based scheme for video watermarking, IEEE, International conference on Communication and Signal Processing, April 3-5, 2013, India.
[29] Paramjit Kaur and Dr. Vijay Laxmi - An Upgraded Approach for Robust Video Watermarking Technique Using Stephens Algorithm, IJCSMC, Vol. 3, Issue. 11, November 2014, pg.612–622
[30] Bhargavi Latha S, Venkata Reddy Dasari and Damodaram Avula - Robust Video Watermarking Using Secret Sharing, SVD, DWT and Chaotic Firefly Algorithm, nternational Journal of Intelligent Engineering and Systems, Vol.11, No.1, 2018
[31] Gopika V Mane, G. G. Chiddarwar, ―Review Paper on Video Watermarking Techniques‖, International Journal of Scientific and Research Publications, Volume 3, Issue 4, April 2013 1 ISSN 2250-3153
[32] R.Lancini, F.Mapelli and S.Tubaro- A robust video watermarking technique in the spatial domain, IEEE, International Symposium on VIPromCom Video/Image Processing and Multimedia Communications 16-19 June 2002.
[33] Amir Houmansadr and Shahrokh Ghaemmaghami- A Novel Video Watermarking Method Using Visual Cryptography, IEEE, International Conference on Engineering of Intelligent Systems
22-23 April 2006.
[34] Chien-Chuan Ko and Bo-Zhi Yang -An Integrated Technique for Video Watermarking., 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007) 11-13 July 2007

Authors



Md. Shahid received his Diploma Engineering in Computer and B.E computer Engineering degree from Jamia Millia Islamia University, New Delhi. He is pursuing M.Tech from Maulana Azad National Urdu University , Gachibowli, Hyderabad, India. His main research area is Digital Image Processing(DIP).



Dr Pradeep Kumar is an Associate Professor in the Department of Computer Science & Information technology at Maulana Azad National Urdu University, Hyderabad (Telangana State). He received his Master's degree in Computer Technology and Applications from Delhi Technological University, formerly Delhi College of Engineering, Delhi University. He completed his Ph.D. from the University School of Information & Communication Technology (USICT), Guru Gobind Singh Indraprastha University (GGSIPU), Delhi. His research interests include software reliability engineering, models for software metrics, machine learning, neural network modeling and soft computing. He has more than 25 publications in journals of international repute including national journals, conferences and proceedings of the international conferences. He is a Member of Association for Computing Machines (ACM) India.

All Rights Reserved © 2018 IJARCET

DroidMark – A Tool for Android Malware Detection using Taint Analysis and Bayesian Network

Dhruv Rathi
Department of Computer Engineering
Delhi Technological University
New Delhi, India
dhruvrathi15@gmail.com

Dr Rajni Jindal
Head of Department, Department of Computer Engineering
Delhi Technological University
New Delhi, India
rajnijindal@dce.ac.in

***Abstract*—** With the increasing user-base of Android devices & advents of technologies such as Internet Banking, delicate user data is prone to be misused by malware and spyware applications. As the app-developer community increases, the quality reassurance could not be justified for every application and a possibility of data leakage arises. In this research, with the aim to ensure the application authenticity, Deep Learning methods and Taint Analysis are deployed on the applications. The detection system named DroidMark looks for possible sinks and sources of data leakage in the application by modelling Android's lifecycles and callbacks, which is done by Reverse Engineering the APK, further monitoring the suspected processes and collecting data in different states of the application. DroidMark is thus designed to extract features from the applications which are fed to a trained Bayesian Network for classification of Malicious and Regular applications. The results indicate a high accuracy of 96.87% and an error rate of 3.13% in detection of Malware in Android devices.

Keywords- *Android; Malware Detection; Bayesian Network; Spyware Detection; Deep Learning; Taint Analysis; Mobile Devices.*

*****

I. INTRODUCTION

Mobile devices with easy access and continuous development on optimizing the system to support multi-functionalities and running high-level applications on relatively lower computationally powered devices have led to high advancements in the devices. Android devices, in particular, are the most popular Operating System(OS) among all with 352 million Smartphones running the OS, which accounts for 81.7% of total industry. The primary reason for the popularity of the Google-owned OS is the huge Open-Source community that is continuously involved in the development of the system. The OS architecture has Linux at the kernel, hence the developers are free to modify the Linux kernel to fit their needs. The ubiquity of the Android Platform has not gone unnoticed by Malware and Spyware developers. Android devices have been targeted in the past by Malicious applications and widespread attacks targeting mass users have been exploited. A malicious application targets the device with the aim to collect user information which otherwise should be kept discreet. Some examples could be credit card information or sharing of location when not needed, sending and intercepting SMS and voice calls, etc.

A plethora of attempts have been made in the past to target the mass user, DroidDream (Android.Rootcager)[1] in 2011, infected 60 different wide used apps, and breached the Android Security sandbox, installed additional malicious software and stealing of data. Android.PjappsM[2], a 2010 botnet infected huge number of devices by that would then launch attacks on websites to steal additional data and infect more devices, initiating a chain reaction.

Malware detection can be modelled using different approaches, as shown in Figure 1[2].

A Signature-based detection technique requires some previous information of the features which are known to be malicious to categorize the program under inspection. In contrast, the Anomaly-based technique uses its knowledge of what is normal to decide the maliciousness of a program under inspection.

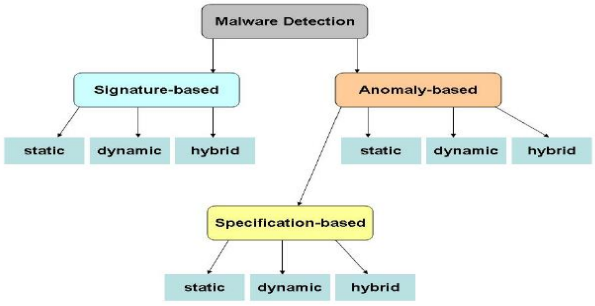


Fig 1. Classification of Malware detection techniques.

The presented research exploits the fact that a malicious application necessarily makes API calls and executes processes which when tested using proper tools proves to be out of context. The research presents a tool DroidMark which maps the context-flow, object-sensitivity, and lifecycle-awareness which results in potential sinks of data leakage, the sinks are tracked by deploying optimized process tracing techniques. The data is collected using various mobile sources, maintained on an allocated server, further data cleaning techniques are employed. The final data of the suspicious processes is obtained and passed through a Deep Learning model of trained Bayesian Network which classifies the application in Regular/Malicious.

II. RELATED WORKS

Utilizing Deep Learning for Android Malware detection has not been explored to a great depth yet, following we have discussed some related works,

Zhenlong Yuan et. al [1] attempts to classify an Android application into Malicious and Normal using Support Vector Machine, a Deep Learning algorithm, achieving a 96.76% detection accuracy on a set of 20,000 Android applications.

Xudong He et. al[2] employs High-Level Petri Nets for modelling and analyzing the Android Permissions Framework and exploring Malicious leakage at the permissions level.

Joshua Abah et. al[3] monitors different API and employs Machine Learning algorithm K Nearest Neighbors to classify the data into Normal or Benign using WEKA tool, reporting an accuracy of 93.75%.

III. EXPERIMENTAL AND COMPUTATIONAL DETAILS

A. Layout

The flow of the tool DroidMark, developed in the research process can be best explained by visualizing it in 3 steps. The first step is called the Taint Analysis. The second step is responsible for the collection of data. The final step is the Deployment of Deep Learning of models to classify the data for detection of Malicious applications.

B. What is Taint Analysis?

Taint Analysis analyzes the application and presents potentially malicious data flow to human analysts or to automated malware detection tools.

FlowDroid: FlowDroid is an open-source tool for static-taint analysis which is integrated into DroidMark to detect sensitive data flows in Android applications. FlowDroid is a full context, flow, field, object sensitive tool with the aim to correctly model the Android lifecycle, UI interaction, system-event handling by analyzing the application's byte-code and configurational files to find potential privacy leaks, either created by mistake or with malicious intentions. In Android app's code there is no Main function unlike normal JAVA code, thus it becomes difficult to obtain the control flow graph simply by reviewing the entry and exit points of the classes, thus as an effective way around the life-cycle functions such as onCreate(), onStart(), onDestroy(), etc. contains various callbacks which notify the application flow, the system events and GUI events.

Targeting the tracking of sensitive data flows and potential Sources and Sinks of data leakage in the Android application, FlowDroid searches for lifecycle and callback methods as well as calls to potential data leakage points. FlowDroid formalized the data flow analysis to a taint analysis on IFDS framework. In short, FlowDroid generates a report which indicates the potential data leakage Sources and Sinks by different API calls and process calls.

Malware used: For the current explanation, the Android-virus Elite[4] is taken, Elite virus forces the following malware into the device:

- Send SMS continuously from the device till the balance is nil.
- Block SMS messenger and related applications.
- Run an unauthorized background process, with a handler on Boot Receiver.
- Wipe out SD-card data completely.
- Retrieve data input by the user in another application.

The dynamic usability of DroidMark allows the system to enter the APK and FlowDroid follows two-step process of analysis consisting of a forward-taint propagation which tracks the tainted value, and a backwards-alias that resolves to alias in the heap model.

IV. DATA LEAKAGE DETECTION: SOURCES AND SINKS

The report generated by FlowDroid using the original features extracted using the data flows and other components contains the list of Source and Sink method names thus generating a feature vector suspicious data leakage points:

```
returnLeakage(application.APK)
{
     result = [(source1_method -> sink1),
        (source1_method -> sink2)
        .
        .
        (sourceM_method -> sinkN_method)]

     return result
}
```

If data flow exists between a source method and a sink method, the feature vector is updated with a value of 1, else it is assigned 0. There are different sources of these sinks such as external libraries which may not be interacting with the application at any points in the lifecycle, thus the feature vector will remain too sparse to produce good results as the number of susceptible sinks are higher than those getting involved.

A. SuSI

To overcome this setback, the SuSI method as discussed in Dil Zhu et. al [4] for enhanced automation of sources and sinks detection from the source code of an Android API. SuSI methodology proposes a categorization into 17 sources and 19 sinks.

S.No.	Source Category	Sink Category
1	LOCATION_INFORMATION	PHONE_CONNECTION
2	NETWORK_INFORMATION	PHONE_CONNECTION
3	FILE_INFORMATION	EMAIL
4	BLUETOOTH_INFORMATION	BLUETOOTH
5	EMAIL	AUDIO
6	UNIQUE_IDENTIFIER	LOCATTON_INFORMATION
7	ACCOUNT_INFORMATION	PHONE_STATE
8	SYNCHRONIZATION_DATA	SYNCHRONIZATION_DATA
9	SMS_MMS	NETWORK
10	SYSTEM_SETTING	SMS_MMS
11	CONTACT_INFORMATION	FILE
12	CALENDAR_INFORMATION	LOG
13	IMAGE	CONTACT_INFORMATION
14	BROWSER_INFORMATION	CALENDAR_INFORMATION
15	NFC	SYSTEM_SETTING
16	DATABASE_INFORMATION	SYNCHRONIZATION_DATA
17	NO_CATEGORY	NFC
18		BROWSER_INFORMATION
19		NO_CATEGORY

Table 1. SuSI categories of sensitive sources and sinks

Providing Elite.APK as the input file to a modified SuSI generates the following outputs: which are suspected to cause data leakage in the malicious application.

Process List: {
'com.samsung.ui',
'datapole.rathi.monitor',
'com.elite.AlarmReceiver',
'com.elite.SMSReceiver', 'com.android.bluetooth',
'android.telephony.SMSManager',
'com.sec.imsservice',
'com.elite.BootReceiver'}

B. Data Collection by Monitoring Sources and Sinks

DroidMark has provided us with a list of suspected processes with the help of FlowDroid and SuSI frameworks.

The next step is the real-time data collection from devices. The monitoring schema is designed based on the fact that a malicious application requires user interaction to activate its functionality on the target device. The analysis of Elite.APK has provided us with, and thus 4 different system process calls, in addition, 1 more feature is added to prepare the feature vectors for the Bayesian Network as indicated in Table-1.

android.telephony.SmsManager
com.elite.BootReceiver
com.elite.SMSReceiver
com.elite.AlarmReceiver
Screen Wake State

Table 2. Monitored processes by DroidMark

Describing the features in Table-2,

A feature table for further reference is created,

feature[] = {android.telephony.SmsManager,
com.elite.BootReceiver,
com.elite.SMSReciever,
com.elite.AlarmReceiver,
ScreenWakeState **}**,

feature[1] controls the incoming and outgoing of SMS services, which is an OS level call, feature[2], feature[3] and feature[4] are packages in Elite.APK which receive SMS, Alarm state Listeners respectively.

feature[4] also logs in whether the device state is of Deep Sleep or is Awake during the Calling and Callback to the listeners of the above processes vector.

A foreground service is created by DroidMark which operates to completely focus on the process in feature[] array, this is done by creating a function killAllProcess(), this function is designed to kill all the running processes which are marked suspected by FlowDroid, which in the present case of Elite.APK are the contents of the array 'feature[]'.

This gives DroidMark, the advantage that the monitoring of features such as SMS Manager will be of the application under process and not that of other parallel running applications.

feature[1]	feature[2]	feature[3]	Wake/Sleep	Type
1	0	1	0	Malicious
1	0	1	1	Normal

Table 3. Classification Example

Considering Table-3, even though the device reports the process of outgoing SMS via the process call android.telephony.SmsManager, this is considered Normal when the device is in awake state(state 1), but is considered Malicious when the device undergoes Deep Sleep (state 1), but is considered Malicious when the device undergoes Deep Sleep (state 0), because the act of sending SMS requires an active user interaction.

Similar feature vectors are parsed and converted into Attribute Relation File Format (ARFF) with the process name and the states appended to each instance and further used as input to train a Bayesian Network.

C. Bayesian Networks

Bayesian networks are a type of Probabilistic Graphical Model. Each node in the graph represents probabilistic dependencies among the corresponding random variables. A Bayesian network B is an annotated acyclic graph that represents a Joint Probability Distribution(JPD) over a set of random variables **V**.

The network is defined by a pair B = <G, Θ>, where G is a Direct Acyclic Graph whose nodes X_1, X_2,...,X_n represents random variables, and whose edges represent the direct dependencies between these variables. The graph G encodes independence assumptions, by which each variable Xi is independent of its non-descendants given its parents in G. The second component Θ denotes the set of parameters of the network. This set contains the parameter $\theta_{xi\,|\pi i}$ = PB(xi|πi) for each realization x_i of X_i conditioned on π_i, the set of parents of Xi in G. Accordingly, B defines a unique Joint Probability Distribution over V, namely:

$$P_B(X_1, X_2,...,Xn) = \prod^{n}_{i=1} P_B(X_i \mid \pi_i) = \prod^{n}_{i=1} \theta_{\mathbf{xi}\,|\boldsymbol{\pi}\mathbf{i}}$$

D. WEKA

Waikato Environment for Knowledge Analysis(WEKA) is an open-source library in JAVA that contains set of algorithms that can be easily implemented in an Android device with a specialized Android application. The Weka.jar file is added as an external library to the Android studio project, into which an appropriate GUI is created to invoke various features programmatically using a set of Java APIs.

V. RESULTS AND DISCUSSION

A. Feeding Data to WEKA

The data collected using different users of the malicious application Elite.APK is collected on a centred Head device which has the DroidMark set-up installed. The data collected is to be fed to the Bayesian Network to complete the Bayesian Network training stage.

The data collected from multiple sources need to be pre-processed in the Attribute Relation File Format (ARFF) format which is the requirement of WEKA.

Sample data collected in ARFF format is explained in Figure 2.

Each attribute has the value 0 or 1, which is decided by whether the system process/ system API is called in the current state of the system by different application packages. The Class attribute is the unknown which needs to be predicted and thus is represented by '?', which is the standard format for unknown variables in ARFF. The @data contains the processes of different applications and data about whether there is a leakage simultaneously activation of any of the data leakage processes, represented by 0/1.

```
@relation RunningProcessVectors
@attribute ProcessName {'com.samsung.ui','datapole.rathi.monitor',
                        'com.elite.AlarmReceiver',
                        'com.elite.SMSReceiver',
                        'com.android.bluetooth',
                        'android.telephony.SMSManager',
                        'com.sec.imsservice','com.elite.BootReceiver'}
@attribute BootReceiver {0,1}
@attribute SMSReceiver {0,1}
@attribute AlarmReceiver {0,1}
@attribute android.telephony.SmsManager {0,1}
@attribute ScreenWake {0,1}
@attribute Class {Regular, Malicious}
@data
'com.samsung.ui',0,0,0,0,1,?
'datapole.rathi.monitor',1,0,1,0,1,?
'com.elite.AlarmReceiver',0,1,1,0,1,?
'com.elite.SMSReceiver',1,1,0,1,1,?
'com.android.bluetooth',0,0,0,0,1,?
'android.telephony.SMSManager',0,1,0,1,1,?
'com.sec.imsservice',0,0,0,0,1,?
'com.elite.BootReceiver',1,1,0,1,1,?
'com.samsung.ui',1,0,1,0,0,?
'datapole.rathi.monitor',1,0,1,0,0,?
'com.elite.AlarmReceiver',1,1,1,0,0,?
'com.elite.SMSReceiver',0,1,0,1,0,?
'com.android.bluetooth',0,0,0,0,0,?
'android.telephony.SMSManager',1,1,0,1,0,?
'com.sec.imsservice',0,0,0,0,0,?
'com.elite.BootReceiver',1,1,1,1,0,?
'com.samsung.ui',1,0,1,0,0,?
'datapole.rathi.monitor',1,0,1,0,0,?
'com.elite.AlarmReceiver',1,1,1,1,0,?
'com.elite.SMSReceiver',1,1,0,1,0,?
'com.android.bluetooth',0,0,0,0,0,?
'android.telephony.SMSManager',1,1,0,1,0,?
```

Fig2. eliteDATA.arff – Parsed data sample in ARFF format

B. *The Criterion for Accuracy of Results*

The output gives us several combinations of criterions to find the accuracy, similar approaches have been used earlier by [13]. In our context, the relevant measures are discussed below,

1. **TPR**:- Rate of True positives (TP), i.e., instances which are correctly classified as a class.
2. **FPR**:- Rate of False Positive (FP), i.e. instances which are falsely classified as a class.
3. **Precision**:- The proportion of instances that are truthful of a class divided by the total instances of that class.
4. **Recall**:- The proportion of instances classified as a given class divided by the actual total in that class**.**
5. **ROC:-** ROC value should be approaching 1 for an efficient classifier.

C. *Output and Accuracy*

Similar data as in Figure 2 is fed in the Bayesian network for the purpose of classifying the application in Malicious/ Normal category.

```
=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D
Relation:     RunningProcessVectors
Instances:    32
Attributes:   7
              ProcessName
              BootReceiver
              SMSReceiver
              AlarmReceiver
              android.telephony.SmsManager
              ScreenWake
              Class
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Bayes Network Classifier
not using ADTree
#attributes=7 #classindex=2
Network structure (nodes followed by parents)
ProcessName(8): SMSReceiver
BootReceiver(2): SMSReceiver
SMSReceiver(2):
AlarmReceiver(2): SMSReceiver
android.telephony.SmsManager(2): SMSReceiver
ScreenWake(2): SMSReceiver
Class(2): SMSReceiver
LogScore Bayes: -172.44387603519843
LogScore BDeu: -218.1484341583698
LogScore MDL: -217.64854929907725
LogScore ENTROPY: -174.32685051408063
LogScore AIC: -199.32685051408063
```

Fig 3. Model structure and SuSI output

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          31               96.875  %
Incorrectly Classified Instances         1                3.125  %
Kappa statistic                          0.9375
Mean absolute error                      0.0678
Root mean squared error                  0.1609
Relative absolute error                 13.4539 %
Root relative squared error             31.9165 %
Total Number of Instances               32
```

Fig 4. Result and Accuracy

```
=== Confusion Matrix ===

  a  b   <-- classified as
 16  0 |  a = 0
  1 15 |  b = 1
```

Fig 5. Confusion Matrix

```
=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.063    0.941      1.000   0.970      0.939    1.000     1.000     0
               0.938    0.000    1.000      0.938   0.968      0.939    1.000     1.000     1
Weighted Avg.  0.969    0.031    0.971      0.969   0.969      0.939    1.000     1.000
```

Fig 6. Summarized Text Results

D. Discussion on Results

From the outputs shown in 3.3, it can be easily inferred that the Bayesian Network provides a very high accuracy, precisely, DroidMark correctly classifies 96.88% of the samples with error percentage of 3.12% as shown in the output of Figure 5, and summarized in Figure 6 in accordance with Criterion for Accuracy as set in 3.2

Discussing the results further in detail in the context of 3.2, the Precision of Malicious and Regular applications is 0.941 and 1.000, and the TPR (which equals the Recall) are 0.938 and 1.000 respectively.

The confusion matrix from Figure 5 shows the correctly classified samples and the misclassified malicious samples from the experiment, it is evident that the malicious class samples classified as being of the Normal class were responsible for the incorrect classification.

An accuracy of 96.88% achieved by the Bayesian Network classifier indicates a significant improvement in detecting a Malicious application, moreover, the time taken to process the collected data is found to be in range of 0.5-10 seconds, thus making the model a good choice for integration in real time devices.

CONCLUSION

In the proposed work, Taint Analysis and Machine Learning approach has been followed to implement an efficient Malware detection system for Android devices. Taint Analysis attempts to construct the life-cycle by Reverse Engineering the APK file of the application under review, this step is performed to find potential sources of sensitive Data Leakage without the consent of the user. Thus, as an output, various processes are marked as Sources and Sinks of leakage. This is done with a modified version of DroidFlow and SuSI frameworks. These processes are monitored for the purpose of data collection via a specialized Android application created with JAVA. WEKA tools were used to train and test a Bayesian Network inside the Android system by feeding the data collected. The results with an accuracy of 96.88% and an error rate of 3.22%, clearly indicate that DroidMark promises an efficient system in a real-time implementation in Android devices,

REFERENCES

[1] Z. Yuan, Y. Lu and Y. Xue, "Droiddetector: android malware characterization and detection using deep learning," in *Tsinghua Science and Technology*, vol. 21, no. 1, pp. 114-123,Feb.2016. doi: 10.1109/TST.2016.7399288

[2] X. He, "Modeling and Analyzing the Android Permission Framework Using High Level Petri Nets," *2017 IEEE International Conference on Software Quality, Reliability and Security (QRS)*, Prague, Czech Republic, 2017, pp. 232-239. doi: 10.1109/QRS.2017.34

[3] Joshua Abah et al. “A MACHINE LEARNING APPROACH TO ANOMALY-BASED DETECTION ON ANDROID PLATFORMS”,International Journal of Network Security & Its Applications (IJNSA) Vol.7, No.6, November 2015

[4] https://github.com/harshalbenake/Android-Elite-Virus/

Accepted Manuscript



Dy^{3+} ions doped oxy-fluoro boro tellurite glasses for the prospective optoelectronic device applications

V. Murali Krishna, Sk. Mahamuda, Rupesh A. Talewar, K. Swapna, M. Venkateswarlu, A.S. Rao

PII: S0925-8388(18)31904-2

DOI: 10.1016/j.jallcom.2018.05.191

Reference: JALCOM 46167

To appear in: *Journal of Alloys and Compounds*

Received Date: 5 February 2018

Revised Date: 15 May 2018

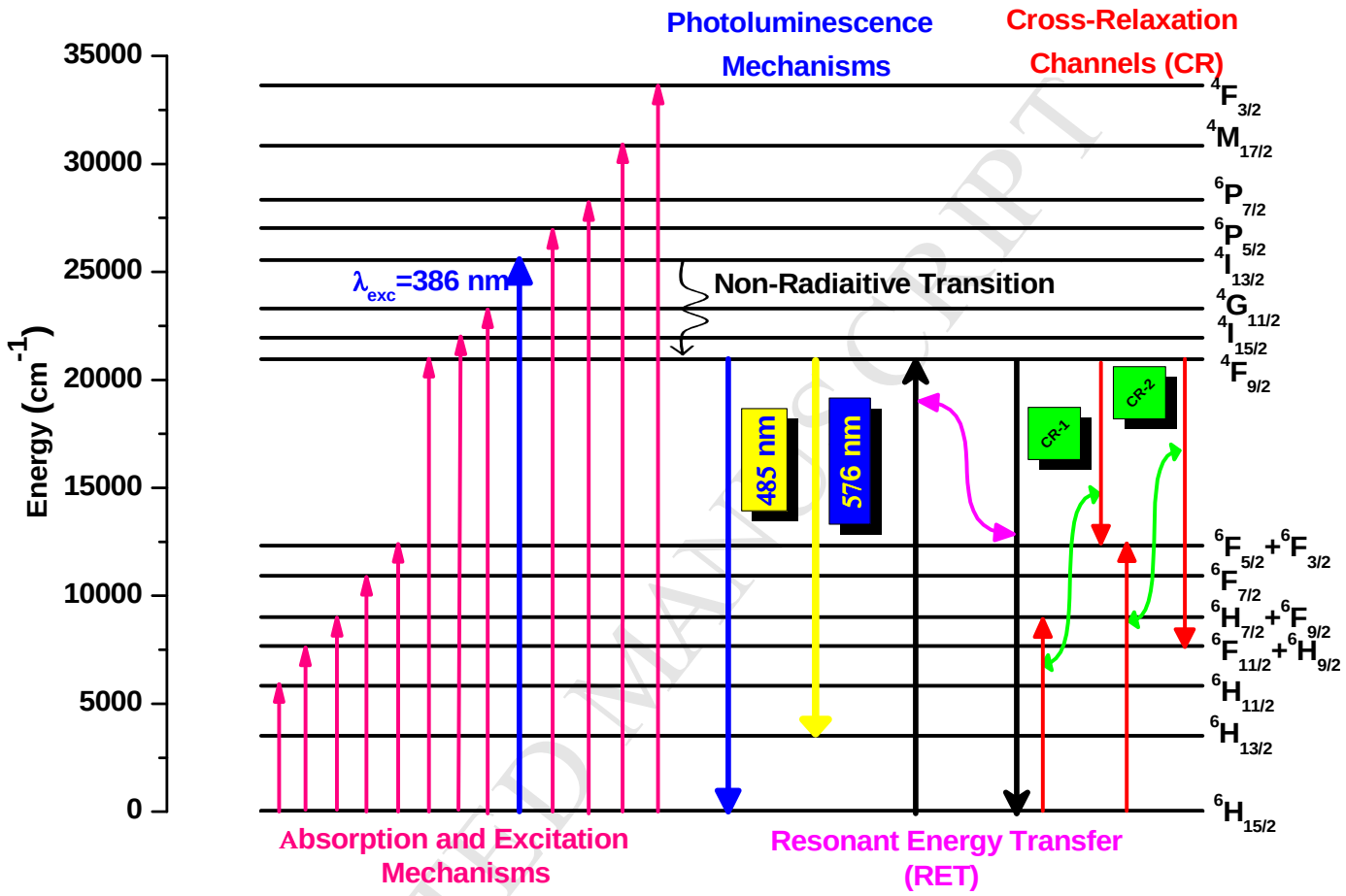
Accepted Date: 17 May 2018

Please cite this article as: V.M. Krishna, S. Mahamuda, R.A. Talewar, K. Swapna, M. Venkateswarlu, A.S. Rao, Dy^{3+} ions doped oxy-fluoro boro tellurite glasses for the prospective optoelectronic device applications, *Journal of Alloys and Compounds* (2018), doi: 10.1016/j.jallcom.2018.05.191.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Graphical Abstract



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Dy^{3+} ions doped oxy-fluoro boro tellurite glasses for the prospective optoelectronic device applications

V. Murali Krishna[1], Sk. Mahamuda[1*], Rupesh A. Talewar[1], K. Swapna[1], M. Venkateswarlu[1] and A. S. Rao[2]

[1]Department of Physics, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur (Dt)-522 502, A.P, India

[2]Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110 042, India

ABSTRACT

Dysprosium (Dy^{3+}) ions doped oxy-fluoro boro tellurite glasses were prepared by using melt quenching technique to study their structural and optical properties using XRD, Raman, absorption, photoluminescence (PL) and PL decay spectral measurements. The XRD and Raman spectral patterns were used to confirm the glassy nature and various functional groups present in the prepared glasses respectively. The oscillator strengths measured from the absorption spectral features were used to evaluate the Judd-Ofelt (J-O) intensity parameters. The PL spectra recorded under 386 nm excitation wavelength show two intense peaks; one in blue (485 nm) and the other one in yellow (576 nm) region corresponding to $^4F_{9/2}\rightarrow{}^6H_{15/2}$ and $^4F_{9/2}\rightarrow{}^6H_{13/2}$ transitions respectively. By correlating the J-O parameters with PL spectral data, various radiative properties like radiative transition probability (A_R), stimulated emission cross section (σ_{se}) and branching ratios (β_R) were evaluated for $^4F_{9/2}\rightarrow{}^6H_{13/2}$ and $^6H_{15/2}$ transitions. Using the PL decay spectral data the experimental lifetimes have been measured and in turn used to measure the quantum efficiency of the prepared glasses. The yellow to blue intensity ratio (Y/B) has been evaluated to understand the white light emission capacity of the host glass by changing the concentration of Dy^{3+} ions. In order to understand white light emission capacity of the prepared glasses, the CIE color co-ordinates have also been evaluated. From all the aforementioned studies, it is concluded that 1.0 mol% of Dy^{3+} ions doped oxy-fluoro boro tellurite glass is aptly suitable for optoelectronic device applications.

Keywords: Dysprosium, Glass, Judd-Ofelt parameters, Photoluminescence, Y/B Ratio

**** Corresponding author,*** **E-mail: mahamuda.ss@gmail.com (Dr. Mahamuda Shaik)**

Tel.: +91-9908452265; Fax: 0863-2388999

ACCEPTED MANUSCRIPT

1. Introduction

Quite recently, solid-state lighting (SSL) technology is budding as the major choice to the present day technologists working in lighting industry to meet the needs of ever increasing energy demands and producing materials for eco friendliness. Hence it became an interesting field of research for the present day scientists/engineers. Especially the research work related to the development of white light emitting diodes (w-LEDs) brought extraordinary advancement in SSL technology because of their extremely good features like low power consumption, high efficiency and longer lifetime (~ 41, 00, 000 hrs) [1-3]. In the construction of w-LEDs, phosphors play an important role but due to several drawbacks the efficiency of w-LEDs decreases [1]. To avoid the problems in the manufacturization of w-LEDs, there is a need for inventing promising materials such as glasses and glass-ceramics to develop cheap, pollution-free and highly efficient w-LEDs [4]. Glasses doped with rare earth (RE) ions play an important role in science and industry due to their excellent chemical, physical and optical properties. Nowadays, development of optical devices doped with RE ions became one of the fascinating fields of research as they are used for the production of devices that are compact as well as efficient. RE doped glasses are widely used as laser materials, optical memory devices, magneto-optical devices, medical lasers, eye safe lasers, flat panel displays, fluorescent lamps, w-LED's etc. because of their ample energy levels due to which luminescence in various regions (visible, NIR and IR) becomes possible [5-14]. Glasses doped with RE ions are also used as optical fibre wave guide lasers and optical amplifiers which provide better opportunities for tuning and Q-switching [15].

Recently, the exploration for new materials increased enormously for RE doped optical devices. One of the key feature of RE doped glasses is that they possess wide transmission ranges from ultraviolet (UV) to mid-infrared (IR); an important factor which influences the quantum yield of the material. Apart from this they have good chemical, mechanical, thermal stabilities, high nonlinear refractive index, and low phonon energies. All the aforementioned novel characteristic features prompted many researchers to study the luminescence properties of glasses doped with RE ions [16]. Recently TeO_2 based glasses have attracted the attention of researchers because of their wide range applications in optical communication system. In general glasses based on TeO_2 possess higher values of refractive index among the oxide glasses. This property of TeO_2 makes them as most promising materials for optoelectronic device applications [17,18]. The third order non-liner optical susceptibility values of tellurium based glasses are also high because of high refractive indices and high polarizability (~ $6.95/10^{-30}$ m^3) of tellurium ions [19]. Tellurite glasses are

ACCEPTED MANUSCRIPT
possessing transparent window between 400 nm to 5 μm, due to which the materials based on tellurites are used in visible to mid-IR applications. Due to their low glass transition temperature, tellurite glasses are also used as optical fibers with good cross-section [20]. Apart from this, these glasses are non-hygroscopic in nature [20]. One more remarkable feature of tellurite glasses is that, when the environment in the network becomes more ionic, the addition of modifiers leads to the formation of trigonal pyramidal [$TeO_{3/2}$] (tp) units instead of trigonal bipyramidal [$TeO_{4/2}$] (tbp) units [21]. Besides having several key features, tellurites cannot form glass alone because of its figure of merit (nearly equal to one). This means that TeO_2 is an intermediate and its glass forming ability depends upon the other species present in the melt whose figure of merit is much less than 1 (Na_2O, B_2O_3 etc.). The TeO_2 glasses have several structural units like TeO_4, TeO_3 and TeO_{3+1} and the bond lengths of Te-O range from 0.185 nm to 0.260 nm and these bonds are very weak due to which tellurite glasses are more open than other oxide glasses [22]. Hence in the present work, TeO_2 is considered in combination with B_2O_3 and Na_2O.

The presence of B_2O_3 in combination with TeO_2 in glasses made the glass structure more complex. In addition, the presence of two glass formers in a glass such as TeO_2 and B_2O_3 can initiate the formation of more mixed structural units as in borosilicate glasses [23]. TeO_2 shows interesting effects on the tetrahedral boron's in the glass network. Several structural studies concluded that glasses with more than 50% B_2O_3 content in combination with TeO_2 shares the vertices of tellurite clusters, isolated BO_4, B_2O_5, BO_3 and condensed borate groups. They also form the continuous network of TeO_4, BO_4 and BO_3 groups and causes phase separation. The binary boro tellurite glass system has a simple eutectic due to which crystalline compounds formation is not possible and the rigidity of the glass structure also increases. Boro tellurites have potential applications particularly in micro-electronics and opto-acoustics because of their peculiar optical and electrical properties [24].

A glass system with good optical transmission window can act as a good optical device. This can be achieved by adding Na_2O to a boro tellurite glass system as stabilizing ion [25]. Addition of Na_2O can reduce the viscosity of the melt and lowers the melting temperature. It also controls the crystallization tendency and makes the integrity of the network more resist [26]. The addition of Na_2O to the glasses not only increases the density and refractive indices of the glasses but also improves the glass forming ability, luminescence intensity, physical characteristics and resistance to moisture. Host glass structure can also be modified by changing the triangular BO_3 structural units to tetrahedral BO_4 units accompanied by the formation of NBO's [27]. To improve the luminescence efficiency of any

ACCEPTED MANUSCRIPT

glass system, the phonon energy of that system must be as low as possible. In general the borate systems possess high phonon energies and such high redundant phonon energies can be effectively reduced by adding fluoride components. The addition of fluorides to a borate system not only decreases the phonon energies but also decreases the dispersion losses because of their optical dispersion and absorption in UV and IR regions. The ionic bonds of the fluorides also influence the refractivity of the glass system [28]. The addition of BaF_2 to a boro tellurite glass system increases the thermal stability because Ba^{2+} ions stabilize the tellurite network by decreasing the network polarizability [16]. It can also decrease the glass transition temperature because it affects the bond energy, electro negativity, number of non-bridging oxygens etc. The glass transition temperature mainly decreases with increase in the number of ionic bonds because; the ionicity of cation-fluorine bonds is more than that of cation-oxygen bonds [29]. The addition of fluorides like BaF_2 and CaF_2 introduces two types of changes in the glass network. One is they strongly bound with the cations in the glass network and the other one is that they can create more number of non-bridging oxygen in the borate network. At high contents of B_2O_3, fluorine enters into the network in the form of B-O-F and increases the activation energy of the system. The addition of BaF_2 and CaF_2 can also builds the glass network with CaF_4 and BaF_4 tetrahedral units and create more free space in the network which in turn more helpful for the incorporation of RE ions into the network [30].

Among various RE ions doped crystalline and amorphous materials, Dy^{3+} ions doped glasses are used as efficient solid state materials in the visible region. These ions are most frequently used as dopants in glasses to study their spectroscopic properties because of their various luminescent channels. The luminescence from the meta-stable state $^4F_{9/2}$ gives higher quantum efficiency values. The pumping is very easy in Dy^{3+} ions doped glasses because they can be easily pumped by blue LED's as well as commercial UV LED's. The visible fluorescence of Dy^{3+} ions contains two intense bands in blue (~ 470-500 nm) and yellow (~ 570-600 nm) regions related to the transitions $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ respectively [31]. The combination of these two transitions gives white light emission. The appropriate white light emission can be achieved by varying the composition of glasses or concentration of Dy^{3+} ions in glasses which is directly related to Y/B intensity ratio. They can also give intense emissions in NIR (1.35 and 3.0 nm) regions. Due to these emissions they possess potential applications in various fields such as lasers, display systems and telecommunications. Further, the fluorescence of Dy^{3+} ion at 1.32 μm originating from $^4F_{11/2} \rightarrow {}^6H_{9/2}$ level is suitable for fibre amplification in optical transmission systems. Recently

ACCEPTED MANUSCRIPT
several researchers have studied the luminescence properties and white light emission capabilities from various hosts [32-34]. Hence, the selection of host materials is very essential for the development of efficient optical devices for white light generation. All the aforementioned diversified scientific patronages offered by various chemicals constituents such as TeO_2, B_2O_3, Na_2O, BaF_2 and CaF_2 prompted us to prepare a glassy system namely oxy-fluoroborate tellurite (OFBT) glasses doped with Dy^{3+} ions to understand their utility in optoelectronic devices such as w-LED's and lasers. A conditional glass former (TeO_2) having many scientific merits added with other species (B_2O_3 and Na_2O) gives a good glass system with wide transparent window. Further addition of alkaline earth fluorides (BaF_2 & CaF_2) to the above species reduces the phonon energies and makes the host matrix more suitable to show relatively better luminescence efficiency when compared with many such host matrices reported in literature.

2. Experimental studies

2.1 Synthesis of glass samples

The glass materials used in the present work have been prepared by using conventional melt quenching technique with the composition $(20\text{-}x)Na_2O\text{-}5BaF_2\text{-}5CaF_2\text{-}60B_2O_3\text{-}10TeO_2\text{-}xDy_2O_3$ (where x = 0.5, 1.0, 1.5, 2.0 and 2.5 mol %). For our convenience these glasses based on Dy^{3+} ion concentration from 0.5 to 2.5 mol% have been labelled as OFBTDy05, OFBTDy10, OFBTDy15, OFBTDy20 and OFBTDy25 respectively. In the present work, analar grade chemicals of B_2O_3 (99.98%), TeO_2 (97%), Na_2O (80%), BaF_2 (99%), CaF_2 (99.5%) and Dy_2O_3 (99.9%) high purity have been considered as the starting materials to prepare the glass samples. All these chemicals have been weighed as per the above mentioned composition and crushed in an agate mortar to get homogeneous mixture. After getting homogeneous mixture, it is poured in a silica crucible and heated at 1100°C in an electric furnace for about 30 min to attain homogeneous melt. The melts were then quenched rapidly on a pre-heated brass mould and pressed quickly by another brass mould. Then the prepared glass samples were annealed at 400°C for 2 hours to eradicate thermal strains/cracks produced inside the glass samples. These glass samples are further polished to get smoothness prior to the spectral measurements.

2.2 Characterization

The X-ray diffraction (XRD) patterns were recorded on XPERT-PRO diffractometer using Cu K_α (λ=1.54059 Å) radiation. The SEM-EDAX of un-doped OFBT glass has been recorded using JEOL JSM-7600F FEG-SEM. The Raman spectra of OFBT glasses have been recorded using BRUKER RFS 27: stand alone FT-Raman spectrometer. The densities and refractive

ACCEPTED MANUSCRIPT
indices for the prepared glass samples have been measured by using the Archimedes principle (water as an immersion liquid) and Brewster's angle method (using 650nm line of He-Ne laser source) respectively. The optical absorption spectra for the prepared glasses were recorded in the spectral wavelength range covering 320-2000 nm with a spectral resolution of 0.1 nm using a JASCO model V-670 UV-vis-NIR spectrophotometer. The excitation, PL and PL decay spectral features were recorded by using a Hitachi-F7000 fluorescence spectrophotometer with 150 Watt xenon lamp as an excitation source. All the measurements were carried out at room temperature.

3. Results and discussion

3.1 XRD spectral analysis

Fig. 1 shows the XRD spectra of an un-doped and Dy^{3+} ions doped OFBT glasses. The spectra do not have any continuous or discrete sharp peaks in the diffraction pattern. The absence of sharp peaks authenticates the amorphous nature of present glasses.

3.2 SEM and elemental analysis

The surface morphology of the un-doped glass has been determined by using a scanning electron microscopy (SEM) and is shown in Fig. 2. The irregular particle shapes and fracture surface shown in Fig. 2 is in consistent with the amorphous nature of the prepared glass material. Fig. 3 represents the EDAX spectrum of the un-doped host glass used for elemental analysis. This spectrum gives the information related to the elements which are present in the host glass. The obtained data confirms the presence of the main constituent's such as boron (B), oxygen (O), fluorine (F), sodium (Na), calcium (Ca) and tellurium (Te) in the host glass. Surprisingly in the spectrum, silicon peak with small intensity has also been observed. This may due to the internal fluorescence peak from Si dead layer of Si-Li detector.

3.3 Raman spectral analysis

The Raman spectrum is an important tool to know the presence of different functional groups in a glassy material. Fig. 4 shows the Raman spectrum of an un-doped OFBT glass recorded in the wavelength range 0-1600 cm^{-1}. However the Raman spectra for Dy^{3+} ions doped OFBT glasses were shown in Fig. 4 as an inset. The Raman spectrum of un-doped OFBT glass and Raman spectra of OFBTDy glasses are showing twelve bands at 81, 460, 780, 917, 984, 1028, 1094, 1217, 1292, 1417, 1485 and 1595 cm^{-1}. No significant shift is observed in Raman peaks with the increase in concentration of Dy^{3+} ions in OFBT glasses. Intensity of Raman band at 780 cm^{-1} decreases in OFBTDy glasses indicating the disruption of six-membered rings containing BO_4 tetrahedra [35]. The assignments of all bands are given in Table 1. The band observed below 200 cm^{-1} is considered as boson peak in which scattering of light takes

ACCEPTED MANUSCRIPT

place due to the disordered structure of glasses [36]. The band at 460 cm^{-1} is attributed to di-borate groups and stretching vibrations of Te-O-Te linkages [37,38]. The band observed at 780 cm^{-1} is very dominant when compared with other Raman bands which is due to the symmetric breathing vibrations of six membered rings with one or two BO_3 triangle replaced by BO_4 tetrahedral and also due to the stretching vibrations of TeO_3 units [37,38]. The band observed at 917 cm^{-1} is due to the vibrations of boroxal rings [38]. The bands observed in the region 980-1220 cm^{-1} are due to the vibrations of tetraborate, pyroborate, diborate and orthoborate groups [37]. The band observed at 1292 cm^{-1} is mainly due to the symmetric stretching vibrations of BO_3 groups [39]. The bands observed in the region 1400-1600 cm^{-1} are ascribed to borate triangles (BO_2-O) linked to tetrahedral units (BO_4) and borate triangles (BO_2-O) linked to other borate triangular units.

3.4 Physical and optical properties

Various physical and optical properties of OFBT glasses have been evaluated taking density and refractive index as input parameters by using the expressions given in literature [40] and are given in Table 2. From Table 2, it is observed that densities and refractive indices of OFBTDy glasses increase with increase in Dy^{3+} ion concentration indicating increase in the non-bridging oxygen (NBOs) in the glass network [5]. Low field strength values represent the high solubility of RE ions in the OFBT glasses [5]. Low values of molecular electronic polarisability show the stability of the OFBTDy glasses. For the present series of glasses the optical basicity values increases with increase in concentration of Dy^{3+} ions which is an indication of increasing covalency in the cation-oxygen bonding [5].

3.5 Absorption spectra & Judd-Ofelt analysis

The absorption spectrum of 1 mol% of Dy^{3+} ions doped OFBT glass in UV-vis-NIR regions is shown in Fig. 5 in the wavelength range 320-2000 nm. The spectrum has eleven absorption peaks, among which six are observed in UV region and five are observed in NIR region. All the absorption bands have been originated from the ground state $^6H_{15/2}$ to various excited states $^6P_{7/2}$, $^6P_{5/2}$, $^4I_{13/2}$, $^4G_{11/2}$, $^4I_{15/2}$, $^4F_{9/2}$, $^6F_{5/2}$, $^6F_{7/2}$, $^6H_{7/2}$, $^6F_{11/2}$ and $^6H_{11/2}$ at wavelengths 347, 362, 387, 425, 452, 473, 800, 896, 1083, 1267 and 1664 nm respectively. All these transitions are assigned on the basis of carnal paper [41]. The absorption bands observed in UV-visible regions are weak (except $^6P_{7/2}$) due to strong absorption of host material in that region. In NIR region, the transition $^6H_{15/2} \rightarrow {}^6F_{11/2}$ at 1267 nm is highly intense and it is considered as hypersensitive transition which obeys the selection rule $|\Delta S| = 0$, $|\Delta L| \leq 2$ and $|\Delta J| \leq 2$. From Fig. 5 it is observed that all the NIR transitions are intense as they are spin allowed ($|\Delta S| = 0$) and orbital allowed ($|\Delta L| = 0$). Due to partially filled f-shells, the

ACCEPTED MANUSCRIPT

nephelauxetic effect will arise, which is a measure of the covalency of rare earth-oxygen bond in the host matrix. As a result, electronic orbitals present in 4f configuration are deformed due to the entering of RE ion in the host matrix. This nephelauxetic effect leads wavelength shift due to the overlap of oxygen orbitals and 4f orbitals which in turn leads to the contraction of energy level structure of the RE ions. To get an idea about the nature of Dy^{3+} ligand bond in the present work, the nephelauxetic ratios (β) and bonding parameters (δ) have been evaluated using the equations available in literature [40] and the bonding parameters are found to be -1.5395, -0.5259, -1.6769, -1.7535 and -2.0220 for OFBTDy05, OFBTDy10, OFBTDy15, OFBTDy20 and OFBTDy25 glasses respectively. The covalent and ionic nature of the ligand field around the RE ions can be decided from the positive and negative values of the bonding parameter. From the obtained values of bonding parameter, it can be seen that the ionic nature dominates in the present series of glasses and it is low for OFBTDy10 glass. The variation of bonding parameter with Dy^{3+} ion concentration is shown in Fig. 6.

The intensities of observed transitions for the present series of glasses have been measured in terms of oscillator strengths using the following expression [1] and are given in Table 3.

$$f_{exp} = 4.318 \times 10^{-9} \int \varepsilon(\vartheta) d\vartheta$$

Here $\varepsilon(\vartheta)$ is the molar extinction coefficient at a wave number ϑ cm^{-1}. The calculated oscillator strengths (f_{cal}) of an electric dipole f-f transition from the ground state, ψ_J, to the excited state, $\psi'_{J'}$ have been measured by using Judd-Ofelt (JO) theory and are given in Table 3.

$$f_{cal(\psi_J, \psi'_{J'})} = \frac{8\pi^2 mc\vartheta}{3h(2J+1)} \times \frac{(n^2+2)^2}{9n} = \sum_{\lambda=2,4,6} \Omega_\lambda (\psi_J \| U^\lambda \|, \psi'_{J'})^2$$

Where c is velocity of light, m is the mass of electron, ϑ is the average energy of the transition (cm^{-1}), J is the total angular momentum of the ground state, h is the plank's constant, n is the refractive index of the host glass sample, Ω_λ (λ = 2, 4, 6) are the J-O intensity parameters and $\|U^\lambda\|^2$ is the doubly reduced matrix element of the unit tensor operator [42].

From the Table 3 it is observed that, OFBTDy10 glass possesses maximum values of oscillator strengths for all the transitions. These variations in oscillator strengths with the variation of Dy^{3+} ions concentration are the direct indication of change in the local structure around the Dy^{3+} ions. Among all the transitions, a transition $^6H_{15/2} \rightarrow {}^6F_{11/2}$ has highest

ACCEPTED MANUSCRIPT

oscillator strength for OFBTDy10 glass. The main reason for the highest oscillator strength of hypersensitive transition is that it possesses large value of $\|U^2\|^2$ matrix element. To explain the environment around the RE ion and its interaction in the glass host, J-O parameters have been evaluated using J-O theory [43,44] and are given Table 4 along with the other reported values [45-54]. From Table 4, it is observed that, the J-O parameters are following the same trend ($\Omega_2 > \Omega_6 > \Omega_4$) in all the glasses. All the glasses possess highest values of Ω_2 parameter which represent the covalency of RE-O bond in the glass matrix. Among all the glasses, OFBTDy10 glass possesses highest values of Ω_2 parameter. Ω_2 is considered as a structural parameter which also depends upon the Dy-O covalency in the surrounding area (ligand field) of Dy^{3+} ions and shows variation in symmetry of the Dy^{3+} ions with its surrounding ligands [55]. This is because of the variation of asymmetry and covalency around the RE ion with change in composition of the glass host [56]. In the present series of glasses Ω_2 values dominates which strongly evident that in the present series of glasses oxygen surroundings around the RE ion is more dominant than the fluorine surroundings [55]. The main cause for the larger magnitude of Ω_2 parameter is the high degree of covalence between Dy^{3+} and oxygen ions and larger disorderliness (low symmetry) of the coordination structure around the Dy^{3+} ion, which was directly associated with the ratio of B/Te content [57]. In the present series of glasses B/Te content ratio is high, due to which number of non-bridging oxygens around the Dy^{3+} ion increases. This in turn leads to the higher disorderliness in the structure by converting the BO_4 units to BO_3 units [57]. Ω_4 and Ω_6 are related to the rigidity and viscosity of the glass host in which RE ions are incorporated [58]. The highest values of Ω_4 and Ω_6 for OFBTDy10 glass show that, this glass possesses rigid structure. These parameters not only influence the physical and chemical properties of the chosen host glass but also strongly affect the transition probabilities because they have the interaction between the central ion and the intermediate environment and crystal-field parameters [59,60].

3.6 Analysis of PL spectra and radiative properties

To analyse the PL spectra and to optimize the best glass, it is very much essential to know the excitation wavelength. Hence the excitation spectrum for OFBTDy10 glass has been recorded at an emission wavelength 576 nm and is shown in Fig. 7. The excitation spectrum consists of eight bands in the region 275-500 nm. The excitation bands are observed at 297, 323, 350, 364, 386, 424, 452 and 472 nm corresponding to the transitions $^6H_{15/2} \rightarrow {}^4F_{3/2}$, $^6H_{15/2} \rightarrow {}^4M_{17/2}$, $^6H_{15/2} \rightarrow {}^6P_{7/2}$, $^6H_{15/2} \rightarrow {}^6P_{5/2}$, $^6H_{15/2} \rightarrow {}^4I_{13/2}$, $^6H_{15/2} \rightarrow {}^4G_{11/2}$, $^6H_{15/2} \rightarrow {}^4I_{15/2}$ and $^6H_{15/2} \rightarrow {}^4F_{9/2}$ respectively. Among all the excitation bands, three bands at 350 nm ($^6H_{15/2} \rightarrow {}^6P_{7/2}$), 364 nm ($^6H_{15/2} \rightarrow {}^6P_{5/2}$) and 386 nm ($^6H_{15/2} \rightarrow {}^4I_{13/2}$) are sharp and intense. Hence, PL spectra for the

ACCEPTED MANUSCRIPT

present series of OFBTDy glasses have been recorded at an excitation wavelength 386 nm. It is well known that the Dy^{3+} ions can also be used as spectroscopic probes to study the local symmetry around the RE ions. For this the phonon side band (PSB) spectrum is very much useful because it shows only the vibrational modes around the RE ion. This PSB can be identified from the excitation spectrum and it originates from RE ions in a matrix which are inhomogeneously broadened because they are surrounded by different host environment. Due to this inhomogeneous distribution, PSB positions are overlapped with the other transitions. The sharp and intense transition $^6H_{15/2} \rightarrow {}^6P_{7/2}$ located at 350 nm is considered as pure electronic transition (PET). A PSB located at 336 nm is strongly related with this PET. By using these PET and PSB transitions, some vital parameters have been estimated like electron-phonon coupling strength 'g' [61] and phonon energy of the glass host 'hω' [62] using the equations given below.

$$g = \frac{\int I_{PSB} d\vartheta}{\int I_{PET}\, d\vartheta} = 56.904 \times 10^{-3}$$

$$h\omega = PSB - PET = 1190\ cm^{-1}$$

The phonon coupling strength value for the present series of glasses is within the range of oxide glasses and the phonon energy of the glass host confirms the high B_2O_3 content in the present borotellurite glass.

Fig. 8 shows the PL spectra of Dy^{3+} ions doped OFBT glasses at an excitation wavelength 386 nm. The PL spectra show two sharp peaks in blue and yellow regions at 485 nm and 576 nm corresponding to the transitions $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ respectively. Among these two transitions, a transition $^4F_{9/2} \rightarrow {}^6H_{15/2}$ in blue region is considered as magnetic dipole transition, whose intensity does not depend upon the local structure of the glass matrix. The other transition $^4F_{9/2} \rightarrow {}^6H_{13/2}$ observed in yellow region is electric dipole transition and its intensity strongly depends upon the local environment of the glass matrix. This transition is considered as hypersensitive transition and it obeys the selection rule $\Delta L=2$ and $\Delta J=2$. The variations of intensities of blue and yellow transitions are shown in Fig. 9. This consideration is valid when the Dy^{3+} ions are situated in low symmetry sites with no inversion centres. Hence the yellow to blue intensity ratio is known as asymmetry ratio of Dy^{3+} ion which varies with the composition or concentration of glass. This Y/B ratio is used as a probe to analyze the structure of the luminescent glass. The variation of the Y/B ratio with the variation of concentration of Dy^{3+} ions is shown in Fig. 10. For the present series of glasses, the Y/B ratio values evaluated from the PL spectra are found to be 0.97, 1.11, 1.02,

ACCEPTED MANUSCRIPT

0.90 and 0.86 for OFBTDy05, OFBTDy10, OFBTDy15, OFBTDy20 and OFBTDy25 glasses respectively. Among all the glasses, the Y/B ratio of OFBTDy10 glass is more than one which is an optimum value to attain white light emission from any luminescent material. Y/B intensity ratio actually depends upon the three structural parameters Ω_2, Ω_4 and Ω_6. The intensity of the yellow emission peak strongly influenced by Ω_2 and Ω_6 parameters because of their square reduced matrix elements (0.0512 and 0.0573). The intensity of blue emission depends upon only Ω_4 parameter (corresponding reduced matrix element - 0.0172) as the remaining reduced matrix elements for this emission are nearly equal to zero. Hence Y/B ratio is proportional to the values of Ω_2 parameter and subsequently it is affected by site asymmetries and electro-negativities of the ligand ions [63].

From Fig. 8 it is observed that the intensity of both blue and yellow luminescent peaks increases with the Dy^{3+} ion concentration up to 1.0 mol% and after which the intensity of both the peaks decreases. Hence concentration quenching occurs after 1.0 mol% of Dy^{3+} ion concentration because of two reasons. One is due to resonant energy transfer (RET) from excited level to ground level of Dy^{3+} ions and other one is due to cross-relaxation (CR) from donor to the acceptor in the glass matrix. Fig. 11 illustrates the partial energy level diagram with all possible mechanisms (absorption, excitation, luminescence, RET and CR channels) in OFBTDy glasses. The possible RET channel ($^4F_{9/2} \rightarrow {}^6H_{15/2} \leftrightarrow {}^6H_{15/2} \rightarrow {}^4F_{9/2}$) and cross-relaxation channels such as CR-1 ($^4F_{9/2} \rightarrow {}^6F_{5/2} \leftrightarrow {}^6H_{15/2} \rightarrow {}^6H_{7/2} + {}^6F_{9/2}$) and CR-2 ($^4F_{9/2} \rightarrow {}^6F_{11/2} + {}^6H_{9/2} \leftrightarrow {}^6H_{15/2} \rightarrow {}^6F_{5/2} + {}^6F_{3/2}$) [64-68] in the Dy^{3+} ions doped OFBT glasses are also mentioned in Fig. 11. Fig. 12 shows the merged portion of the excitation and emission peaks in which possible RET can occur among the Dy^{3+} ions.

In order to know the best suitable material for laser and optoelectronic device, several radiative properties like radiative transition probabilities (A_R), radiative branching ratios (β_R) and radiative lifetimes (τ_R) have been calculated from J-O parameters [69] and are given in Table 5.

The radiative transition probabilities $A_R\left(\psi_J, \psi'_{J'}\right)$ have been determined by using the following equation.

$$A_R(\Psi_J, \Psi'_{J'}) = \frac{64\pi^4\vartheta^3}{3h(2J+1)}\left[\frac{(n^2+2)^2}{9n} S_{ED}(\Psi_J, \Psi'_{J'}) + n\, S_{MD}(\Psi_J, \Psi'_{J'})\right]$$

Here A_R is the spontaneous emission probability per unit time, Ψ_J, $\Psi'_{J'}$ are the initial and final states, ϑ (cm^{-1}) is the energy gap between the states, n is the refractive index of the medium, S_{ED} and S_{MD} are the electric and magnetic dipole line strengths respectively

ACCEPTED MANUSCRIPT

The total radiative transition probability (A_T) of an excited state is the sum of the terms calculated over all the terminal states

$$A_T(\Psi_J, \Psi'_{J'}) = \sum_{\Psi'_{J'}} A_R(\Psi_J, \Psi'_{J'})$$

A_T is related to the radiative lifetime (τ_R) of an excited state and is expressed by using the following expression

$$\tau_R(\Psi_J, \Psi'_{J'}) = \frac{1}{A_T(\Psi_J, \Psi'_{J'})}$$

The fluorescence branching ratio from the relaxing state to a particular lower lying state is given by

$$\beta_R(\Psi_J, \Psi'_{J'}) = \frac{A_R(\Psi_J, \Psi'_{J'})}{\sum_{\Psi'_{J'}} A_R(\Psi_J, \Psi'_{J'})}$$

The branching ratios can be used to predict the relative intensities of all emission lines originating from a given excited state ($\Psi'_{J'}$). The experimental branching ratio (β_{exp}) can be found from the relative areas of the emission bands.

The peak position (λ_p), effective bandwidths (λ_{eff}) and stimulated emission cross-section (σ_{se}) were evaluated from the PL spectra using the following equations and are given in Table 6. The peak stimulated emission cross-section, σ_{se} (Ψ_J, $\Psi'_{J'}$) between the states (Ψ_J) and ($\Psi'_{J'}$) having a probability of $A_R(\Psi_J, \Psi'_{J'})$ has been evaluated using the following expression

$$\sigma_{se}(\Psi_J, \Psi'_{J'}) = \frac{{\lambda_p}^4}{8\pi c n^2 \Delta\lambda_{eff}} A_R(\Psi_J, \Psi'_{J'})$$

Here λ_p is the transition peak wavelength n is the refractive index and $\lambda_{eff}$is its effective linewidth found by dividing the area of the emission band by its average height which is represented as follows.

$$\Delta\lambda_{eff} = \int \frac{I(\lambda)}{I_{max}} d\lambda$$

Here $I(\lambda)$ is the intensity at the wavelength λ and I_{max} is the maximum emission intensity. Good laser transitions are characterized by large cross-sections for stimulated emission. For the present series of OFBTDy glasses, the magnitude of β_R for the $^4F_{9/2} \rightarrow {}^6H_{13/2}$ (Yellow) lasing transition was found to be more when compared with other transition $^4F_{9/2} \rightarrow {}^6H_{15/2}$ (Blue). Moreover, the radiative and experimental branching ratios (β_R & β_{exp}) for the $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transition are found to be more than 0.50 which is an indication that the present series of glasses can be used for efficient yellow laser emission. Among all the glasses, OFBTDy10 glass possesses highest magnitude of β_R & β_{exp} values. Hence this glass can be

ACCEPTED MANUSCRIPT

suggested as a best glass for efficient yellow laser emission. Furthermore, this glass also possesses highest values of σ_{se}, $\sigma_{se}x\tau_R$ which represents the possibility to achieve high gain for good lasers. Hence OFBTDy10 glass is apt for visible laser applications in yellow region.

3.7 PL decay spectral analysis

The conclusions regarding the nature of quenching mechanism can be drawn from the concentration dependence of the PL decay rates. PL decay measurements of different concentrations of Dy^{3+} ions doped OFBT glasses were performed and systemically investigated. Fig. 13 and 14 show the PL decay curves of $^4F_{9/2}\rightarrow{}^6H_{13/2}$ (576 nm) emission transition for all OFBT glasses under 386 nm excitation. For low concentration of Dy^{3+} ions in OFBT glass, the decay curve exhibits a nearly single exponential decay with a decay time of 1.75 ms (Fig. 14 curve a). The lifetime in the order of millisecond is one of the characteristics of the f→f transitions of dysprosium ions [70]. It can be seen that with increase in Dy^{3+} ion concentration, the decay curve deviates from single exponential to non-exponential. This non-exponential behaviour arises due to the energy transfer among Dy^{3+}-Dy^{3+} ions [71]. By carefully observing the decay curves, it can be seen that the non-exponential curves are decomposed into two exponential components (Fig. 14 b & c). The slow component of decay time 1.75 ms (Fig. 14 curve c) corresponds to Dy^{3+} ions away from Dy^{3+}. The second component is much faster with a decay time of 1.09 ms (Fig. 14 b) corresponds to Dy^{3+} ions close to Dy^{3+}. With increasing Dy^{3+} concentration, the faster component becomes dominant and for 2.5 mol% Dy^{3+}, the slower component is quite subdued.

The average decay time (τ_{avg}), which is also called as experimental lifetime (τ_{exp}) has been evaluated from the recorded decay profiles using the following expression [72,73]

$$I = I_0 + a_1 * \exp(-t/\tau_1) + a_2 * \exp(-t/\tau_2)$$

Here, I_0 is the intensity at t=0, and I is intensity at time t, a_1 and a_2 are the amplitudes of decay constants and τ_1 and τ_2 are the luminescence lifetimes for fast and slow channels of decay respectively. The average value of decay time (τ_{avg}) also called as experimental lifetime (τ_{exp}) is calculated by using the formula

$$\tau_{avg} = \frac{a_1\tau_1^2 + a_2\tau_2^2}{a_1\tau_1 + a_2\tau_2}$$

The τ_{exp} measured from the PL decay spectral profiles are given in Table 7 along with radiative lifetimes (τ_R). The τ_{exp} values are found to be decreasing with the increase in Dy^{3+} ion concentration in OFBT glasses. This is an indicative representation of non-radiative energy transfer from donor to acceptor ions.

ACCEPTED MANUSCRIPT

The non-radiative (W_{NR}) energy transfer has been estimated using the following expression

$$\frac{1}{\tau_{exp}} = \frac{1}{\tau_R} + W_{MP} + W_{ET}$$

Here W_{MP} is multiphoton relaxation rate and W_{ET} is the cross-relaxation rate. In Dy^{3+} ions doped materials, W_{MP} values are negligible because of large gap ~7400 cm^{-1} between $^4F_{9/2}$ to $^4F_{1/2}$ levels which is very much higher than the phonon energy of present host glass. Hence in the present series of glasses, the non-radiative relaxation (NR) takes place due to energy transfer and cross-relaxation. The W_{ET} values are measured using the following expression and are given in Table 7.

$$W_{NR} = W_{ET} = \frac{1}{\tau_{exp}} - \frac{1}{\tau_R}$$

Among all the glasses, OFBTDy10 glass has lowest values of non-radiative energy transfer rates which lead to highest quantum efficiency. The quantum efficiency of Dy^{3+} ions doped OFBT glasses have been evaluated using the following expression and are given in Table 7 [74].

$$\eta = \frac{\tau_{exp}}{\tau_R} \text{ x } 100$$

From Table 7 it is observed that, OFBTDy10 glass possesses maximum quantum efficiency which is a key factor for a luminescent material. From all the evaluated parameters such as branching ratios, stimulated emission cross-sections and quantum efficiencies, it can be suggested that OFBTDy10 glass is aptly suitable as a luminescent material to give yellow emission at 576 nm. The variations of all the above mentioned parameters with Dy^{3+} ion concentration is also shown in Fig. 15.

3.8 CIE color co-ordinates evaluation and importance of color correlated temperature (CCT)

The commission international de i'eclairage (CIE) 1931 has been used to calculate the chromaticity coordinates that are useful in assessing the performance of the Dy^{3+} ions doped glasses for white light emission. The CIE color coordinates (x, y) evaluated using PL spectral data for the prepared glasses are given in Table 8 [75]. The CIE color chromaticity diagram for all the glasses is shown in Fig. 16. From Fig. 16, it is observed that with the increase in Dy^{3+} ion concentration, the CIE color coordinates are found to be shifted towards slightly greenish yellow region. Among all the glasses, OFBTDy10 glass color coordinates (x, y) are

ACCEPTED MANUSCRIPT

very close to white light region. It is known that the perfect white light emission from Dy^{3+} ions doped materials can be achieved by adjusting the yellow to blue intensity ratio; the Y/B ratio in turn can be changed by changing the Dy^{3+} ion concentration.

The quality of the white light emission from the luminescent material can also be assessed using CCT [76]. This CCT provides the information of a closest plankian black body radiation to the operating point on the CIE color chromaticity diagram. In the present series of glasses, OFBTDy10 glass coordinates are near to the white light emission region and its CCT value is found to be 5328 K. This CCT of OFBTDy10 glass is nearer to day light (5500 K). Hence OFBTDy10 glass can be used as a warm white light source. The color purity of the OFBTDy10 glass has also been estimated using the expression given in literature [76]. The color purity values must be as low as possible for getting white light emission. For OFBTDy10 glass, the color purity value is 6.143×10^{-2}. It is nearly equal to some reported glasses and higher than YAG:Dy^{3+} phosphor and NbFSDy0.1 glass [77]. All the aforementioned results finally suggest the utility of OFBTDy10 glass for the development of laser illumination and as well as w-LEDs. The comparison of results with the other reported values are also given in Table 9 [12,77-80].

4. Conclusions

Different concentrations of Dy^{3+} ions doped OFBT glasses have been prepared by using melt quenching technique. The structural and optical properties of the prepared glasses have been studied through XRD, Raman, absorption, PL and PL decay spectral measurements. The amorphous nature and various functional groups present in the prepared glasses have been studied by XRD and Raman spectral studies respectively. The bonding parameters evaluated from the absorption spectral data reveals that the nature of bonding between Dy^{3+} ions and oxygen ligands is covalent. This result is also in consistent with the J-O parameters derived from the oscillator strengths of the absorption spectral features. From the excitation spectrum electron-phonon coupling strength and phonon energy of the glass host have been evaluated. From the PL spectra various radiative properties have been evaluated for the two intense peaks observed in blue (485 nm) and yellow (576 nm) regions corresponding to the transitions $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ respectively. In the present series of glasses, the experimental lifetime's decreases with the increase in Dy^{3+} ion concentration due to energy transfer between Dy^{3+} ions. Using experimental lifetimes, quantum efficiency of the present series of glasses has been measured. From the evaluated branching ratios, stimulated emission cross-sections and quantum efficiency, it is suggested that the OFBTDy10 glass is quite suitable as a luminescent material for visible laser applications. From the evaluated CIE

ACCEPTED MANUSCRIPT

colour coordinates, CCT and colour purity parameters, it is concluded that the OFBT glass with 1.0 mol% of Dy^{3+} ions is aptly suitable for white light generation and is useful in SSL technology.

Acknowledgement

The authors, Prof. A.S. Rao (File Number: EMR/2016/007766), Dr. Sk. Mahamuda (File Number: ECR/2016/000376) and Dr. K. Swapna (File Number: ECR/2015/000335) are thankful to Department of Science and Technology (DST), Govt. of India, New Delhi for the award of major research projects to them.

ACCEPTED MANUSCRIPT

References

[1] Sk. Mahamuda, K. Swapna, P. Packiyaraj, A. Srinivasa Rao, G. Vijaya Prakash, Lasing potentialities and white light generation capabilities of Dy^{3+} doped oxy-fluoroborate glasses, J. Lumin. 153 (2014) 382-392.

[2] Nisha Deopa, A.S. Rao, Ankur Choudhary, Shubham Saini, Abhishek Navhal, M. Jayasimhadri, D. Haranath, G. Vijaya Prakash, Photoluminescence investigations on Sm^{3+} ions doped borate glasses for tricolor w-LEDs and lasers, Mater. Res. Bull. 100 (2018) 206-212.

[3] K. Swapna , Sk. Mahamuda, A. Srinivasa Rao, M. Jayasimhadri , Suman Shakya , G. Vijaya Prakash, Tb^{3+} doped Zinc Alumino Bismuth Borate glasses for green emitting luminescent devices, J. Lumin. 156 (2014) 180-187.

[4] Ch. B. Annapurna Devi, Sk. Mahamuda, M. Venkateswarlu, K. Swapna, A. Srinivasa Rao, G. Vijaya Prakash, Dy^{3+} ions doped single and mixed alkali fluoro tungsten tellurite glasses for LASER and white LED applications, Opt. Mater. 62 (2016) 569-577.

[5] Sk. Mahamuda, K. Swapna, M. Venkateswarlu, A. Srinivasa Rao, Suman Shakya, G. Vijaya Prakash, Spectral characterization of Sm^{3+} ions doped Oxy-fluoroborate glasses for visible orange luminescent applications, J. Lumin. 154 (2014) 410-424.

[6] Ritu Sharma, A.S. Rao, Nisha Deopa, M. Venkateswarlu, M. Jayasimhadri, D. Haranath, G. Vijaya Prakash, Spectroscopic study of Pr^{3+} ions doped Zinc Lead Tungsten Tellurite glasses for visible photonic device applications, Opt. Mater. 78 (2018) 457-464.

[7] Rupesh A. Talewar, Sk. Mahamuda, K. Swapna, M. Venkateswarlu, A.S. Rao, Spectroscopic studies of Sm^{3+} ions doped alkaline-earth chloro borate glasses for visible photonic applications, Mater. Res. Bull. 105 (2018) 45-54.

[8] Nisha Deopa, A.S. Rao, Spectroscopic studies of single near ultraviolet pumped Tb^{3+} doped Lithium Lead Alumino Borate glasses for green lasers and tricolour w-LEDs, J. Lumin. 194 (2018) 56-63.

[9] Ruiyu Mi, Jian Chen, Yangai Liu, Lefu Mei, Jiangyan Yuan, Yufei Xia, Zhaohui Huang, Minghao Fang, Crystal structure and luminescence property of a novel single-phase white light emission phosphor $KCaBi(PO_4)_2:Dy^{3+}$, Mater. Res. Bull. 86 (2017) 146-152.

[10] Mingye Ding, Yuting Li, Daqin Chen, Hongwei Lu, Junhua Xi, Zhenguo Ji, Hexagonal crown-capped $NaYF_4:Ce^{3+}/Gd^{3+}/Dy^{3+}$ microrods: Formation mechanism, energy transfer and luminescence properties, J. Alloys Compd. 658 (2016) 952-960.

ACCEPTED MANUSCRIPT

[11] Mingye Ding, Min Xu, Daqin Chen, A new non-contact self-calibrated optical thermometer based on Ce^{3+}→Tb^{3+}→Eu^{3+} energy transfer process, J. Alloys Compd. 713 (2017) 236-247.

[12] C.R. Kesavulu, H.J. Kim, S.W. Lee, J. Kaewkhao, N. Chanthima, Y. Tariwong, Physical, vibrational, optical and luminescence investigations of Dy^{3+} doped yttrium calcium silicoborate glasses for cool white LED applications, J. Alloys Compd. 726 (2017) 1062-1071.

[13] M. Vijayakumar, K. Viswanathan, K. Marimuthu, Structural and optical studies on Dy^{3+}:Tb^{3+} co-doped zinc lead fluoro-boro phosphate glasses for white light applications, J. Alloys Compd. 745 (2018) 306-318

[14] C. Gheorghe, S. Hau, L. Gheorghe, F. Voicu, M. Greculeasa, M. Enculescu, K.N. Belikov, E. Yu. Bryleva, O.V. Gaiduk, Yellow laser potential of cubic $Ca_3(Nb,Ga)_5O_{12}$:Dy^{3+} and $Ca_3(Li,Nb,Ga)_5O_{12}$:Dy^{3+} single crystals, J. Alloys Compd. 739 (2018) 806-816.

[15] K. Swapna, Sk. Mahamuda, A. Srinivasa Rao, M. Jayasimhadri, T. Sasikala, L. Rama Moorthy, Visible fluorescence characteristics of Dy^{3+} doped zinc alumino bismuth borate glasses for optoelectronic devices, Ceram. Int. 39 (2013) 8459-8465.

[16] Agnieszka Marczewska, Marcin Sroda, Marek Nocun, Thermal and spectroscopic characterization of gallium-tellurite glasses doped BaF_2 and PbO, J. Non-Cryst. Solids 464 (2017) 104-114.

[17] N.V. Ovcharenko, T.V. Smirnowa, High refractive index and magneto–optical glasses in the systems TeO_2–WO_3–Bi_2O_3 and TeO_2–WO_3–PbO, J. Non-Cryst. Solids 291 (2001) 121-126.

[18] P. Hagenmuller [Ed.], Inorganic Solid Fluorides, Chemistry and Physics, Academic Press, New York 1985.

[19] S. Fujino and K. Morinaga. Material dispersion and its compositional parameter of oxide glasses, J. Non-Cryst. Solids 222 (1997) 316-320.

[20] W. Widanarto, M.R. Sahar, S.K. Ghoshal, R. Arifin, M.S. Rohani, K. Hamzah, M. Jandra, Mater. Chem. Phys. 138 (2013) 174-178.

[21] Nirmal Kaur, Atul Khanna, Structural characterization of borotellurite and aluminoborotellurite glasses, J. Non-Cryst. Solids 404 (2014) 116-123.

[22] J.C. Knight, Photonic crystal fibers, Nature 424 (2001) 847-851.

[23] Tetsuaki Nishida, Miho Yamada, Hiroshi Ide, Y. Takashima, Correlation between the structure and glass transition temperature of potassium, magnesium and barium tellurite glasses, J. Mater. Sci. 25 (1990) 3546-3550.

ACCEPTED MANUSCRIPT

[24] Yang Yanmin, Liu Yanzhou, Cai Peiqing, Ramzi Maalej, Hyo Jin Seo, Thermal stability and spectroscopic properties of Ho^{3+} doped tellurite-borate glasses, J. Rare Earths 33 (2015) 939-945.

[25] J.C. Mclaughlin, S.L. Tagg, J.W. Zwanziger, D.R. Haeffner, S.D. Shastri, The Structure of tellurite glass: a combined NMR, neutron diffraction and X-ray diffraction study, J. Non-Cryst. Solids 274 (2000) 1-8.

[26] E.A. Abou Neel, D.M. Pickup, S.P. Valappil, R.J. Newport, J.C. Knowles, Bioactive functional materials: a perspective on phosphate-based glasses, J. Mater. Chem. 19 (2009) 690-701.

[27] L. Shamshad, G. Rooh, K. Kirdsiri, N. Srisittipokakun, B. Damdee, H.J. Kim, J. Kaewkhao, Effect of alkaline earth oxides on the physical and spectroscopic properties of Dy^{3+} doped Li_2O-B_2O_3 glasses for white emitting material application, Opt. Mater. 64 (2017) 268-275.

[28] Fuxi Gan, Optical properties of fluoride glasses: a review, J. Non-Cryst. Solids 184 (1995) 9-20.

[29] T. Konishi, T. Sakamoto, T. Kikkojin, A. Yasumori, S. Suehara, S. Todoroki , S. Inoue, Combinatorial approach to new glasses, Appl. Surf. Sci. 252 (2006) 2450-2455.

[30] H. Doweidar, K. El-Egili, A. Altawaf, Structural units and properties of BaF_2–PbF_2–B_2O_3 glasses, J. Non-Cryst. Solids 464 (2017) 73-80.

[31] Chenxiao Hua, Lifan Shen, Edwin Yue Bun Pun, Desheng Li, Hai Lin, Dy^{3+} doped tellurite glasses containing silver nanoparticles for lighting devices, Opt. Mater. 78 (2018) 72-81.

[32] P.P. Pawar, S.R. Munishwar, S. Gautam, R.S. Gedam, Physical, Thermal, Structural and Optical Properties of Dy^{3+} doped Lithium Alumino-borate Glasses for Bright W-LED, J. Lumin. 183 (2017) 79-88.

[33] Shihua Liu, Jianyue He, Zhaoyi Wu, Jung Hyun Jeong, Bin Deng, RuijinYu, Preparation and study on the spectral properties of garnet-type $Li_3Gd_3Te_2O_{12}$:Dy^{3+} single-phase full-color phosphor, J. Lumin. 200 (2018) 164-168.

[34] P.P. Pawar, S.R. Munishwar, R.S. Gedam, Intense white light luminescent Dy^{3+} doped lithium borate glasses for W-LED: A correlation between physical, thermal, structural and optical properties, Solid State Sci. 64 (2017) 41-50.

[35] B.N. Meera and J. Ramakrishna, Raman spectral studies of borate glasses, J. Non-Cryst. Solids 159 (1993) 1-21

ACCEPTED MANUSCRIPT

[36] Mahesh M. Hivrekara, D.B. Sable, M.B. Solunke, K.M. Jadhav, Network structure analysis of modifier CdO doped sodium borate glass using FTIR and Raman spectroscopy, J. Non-Cryst. Solids 474 (2017) 58-65.

[37] Rawan El Hayek, Frédérique Ferey, Pierre Florian, Alexander Pisch, Daniel R. Neuville, Structure and properties of lime alumino-borate glasses, Chem. Geol. 461 (2017) 75-81.

[38] Nisha Deopa, A.S. Rao, Sk. Mahamuda, Mohini Gupta, M. Jayasimhadri, D. Haranath, G. Vijaya Prakash, Spectroscopic studies of Pr^{3+} doped lithium lead alumino borate glasses for visible reddish orange luminescent device applications, J. Alloys Compd. 708 (2017) 911-921.

[39] Amarjot Kaur, Atul Khanna, Lyubomir I. Aleksandrov, Structural, thermal, optical and photo-luminescent properties of barium tellurite glasses doped with rare-earth ions, J. Non-Cryst. Solids 476 (2017) 67-74.

[40] A.S. Rao, Y.N. Ahammed, R.R. Reddy, T.V.R. Rao, Spectroscopic studies of Nd^{3+}-doped alkali fluoroboro phosphate glasses, Opt. Mater. 10 (1998) 245-252.

[41] W.T. Carnall, P.R. Fields, K. Rajnak, Electronic Energy Levels in the Trivalent Lanthanide Aquo Ions. I. Pr^{3+}, Nd^{3+}, Pm^{3+}, Sm^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, and Tm^{3+}, J. Chem. Phys. 49 (1968) 4424-4443.

[42] C.K. Jayasankar, E. Rukmini, Optical properties of Sm^{3+} ions in zinc and alkali zinc borosulphate glasses, Opt. Mater. 8 (1997) 193-208.

[43] B.R. Judd, Optical absorption intensities of rare-earth ions, Phys. Rev. 127 (1962) 750-761.

[44] G.S. Ofelt, Intensities of crystal spectra of rare-earth ions, J. Chem. Phys. 37 (1962) 511-520.

[45] P. Suthanthirakumar, K. Marimuthu, Investigations on spectroscopic properties of Dy^{3+} doped zinc telluro-fluoroborate glasses for laser and white LED applications, J. Mol. Struct. 1125 (2016) 443-452.

[46] N. Vijaya, K. Upendra Kumar, C.K. Jayasankar, Dy^{3+} doped zinc fluorophosphate glasses for white luminescence applications, Spectrochim. Acta A 113 (2013) 145-153.

[47] J. Pisarska, Optical properties of lead borate glasses containing Dy^{3+} ions, J. Phys. Condens. Matter 21 (2009) 285101.

[48] D.V.R. Murthy, B.C. Jamalaiah, A. Mohan Babu, T. Sasikala, L. Rama Moorthy, The luminescence properties of Dy^{3+}-doped alkaline earth titanium phosphate glasses, Opt. Mater. 32 (2010) 1112-1116.

ACCEPTED MANUSCRIPT

[49] R.T. Karunakaran, K. Marimuthu, S. Surendra Babu, S. Armugam, Structural, optical and thermal investigations on Dy^{3+} doped $NaF–Li_2O–B_2O_3$ glasses, Phys. B: Condens. Matter 404 (2009) 3995-4000.

[50] R. Vijaya Kumar, G. Venkataiah, K. Marimurthu, White light simulation and luminescence studies on Dy^{3+} doped Zinc borophosphate glasses, Phys. B: Condens. Matter 457 (2015) 287-295.

[51] C.R. Kesavulu, C.K. Jayasankar, White light emission in Dy^{3+}-doped lead fluoro phosphate glasses, Mater. Chem. Phys. 130 (2011) 1078-1085.

[52] K.V. Krishnaiah, K. Upendra Kumar, C.K. Jayasankar, Spectroscopic properties of Dy^{3+}-doped oxyfluoride glasses for white light emitting diodes, Mater. Exp. 3 (2013) 61-70.

[53] K.K. Mahato, Anitha Rai, S.B. Rai, Optical properties of Dy^{3+} doped in oxyfluoroborate glass, Spectrochim. Acta part A 61 (2005) 431-436.

[54] B. Shanmugavelu, V.V. Ravi Kanth Kumar, Luminescence studies of Dy^{3+} doped bismuth zinc borate glasses, J. Lumin. 146 (2014) 358-363.

[55] G. Wang, S. Dai, J. Zhang, L. Wen, J. Yang, Z. Jiang, Intense up-conversion luminescence and effect of local environment for Tm^{3+}/Yb^{3+} co-doped novel $TeO_2–BiCl_3$ glass system, Spectrochim. Acta Part A 64 (2006) 349–354.

[56] X. Zou, H. Toratani, Evaluation of spectroscopic properties of Yb^{3+}-doped glasses, Phys. Rev. B 52 (1995) 15889–15897.

[57] Vu Phi Tuyen, Bounyavong Sengthong, Vu Xuan Quang, Phan Van Do, Ho Van Tuyen, Le Xuan Hung, Nguyen Trong Thanh, Masayuki Nogami, Tomokatsu Hayakawa, Bui The Huy, Dy^{3+} ions as optical probes for studying structure of boro-tellurite glasses, J. Lumin. 178 (2016) 27-33.

[58] F. Auzel, f–f oscillator strengths, hypersensitivity, branching ratios and quantum efficiencies discussed in the light of forgotten results, J. Alloys Compd. 380 (2004) 9-14.

[59] Raouf El-Mallawany, Tellurite glasses Part 1. Elastic properties, Mater. Chem. Phys. 53 (1998) 93-120.

[60] H. Gebavi, D. Milanese, R. Balda, M.Ivanda, F. Auzel, J. Lousteau, J. Fernandez, M. Ferraris, Novel Tm^{3+} doped fluorotellurite glasses with enhanced quantum efficiency, Opt. Mater. 33 (2011) 428-437.

[61] S. Todoroki, S. Tanabe, K. Hirao, N. Soga, Phonon sideband spectra and local structure around Eu^{3+} ions in sodium silicate glasses, J. Non-Cryst. Solids 136 (1991) 213-218.

ACCEPTED MANUSCRIPT

[62] K. Selvaraju, K. Marimuthu, T.K. Seshagiri, S.V. Godbole, Thermal, structural and spectroscopic investigations on Eu^{3+} doped boro-tellurite glasses, Mater. Chem. Phys. 131 (2011) 204-210.

[63] P. Van Do, V.P. Tuyen, V.X. Quang, N.T. Thanh, V.T.T. Ha, H. Van Tuyen, N.M. Khaidukov, J. Marcazzó, Y.I. Lee, B.T. Huy, Optical properties and Judd-Ofelt parameters of Dy^{3+} doped K_2GdF_5 single crystal, Opt. Mater. 35 (2013) 1636-1641.

[64] J. Suresh Kumar, K. Pavani, A. Mohan Babu, Neeraj Kumar Giri, S.B. Rai, L. Rama Moorthy, Fluorescence characteristics of Dy^{3+} ions in calcium fluoroborate glasses, J. Lumin. 130 (2010) 1916-1923.

[65] T. Srihari, C.K. Jayasankar, Fluorescence properties and white light generation from Dy^{3+} doped niobium phosphate glasses, Opt. Mater. 69 (2017) 87-95.

[66] P. Karthikeyan, R. Vijayakumar, K. Marimuthu, Luminescence studies on Dy^{3+} doped Calcium boro-tellurite glasses for White light applications, Phys. B: Condens. Matter 521 (2017) 347-354.

[67] F. Zaman, J. Kaewkhao, N. Srisittipokakun, N. Wantana, H.J. Kim, G. Rooh, Investigation of luminescence and laser transition of Dy^{3+} in Li_2O-Gd_2O_3-Bi_2O_3-B_2O_3 glasses, Opt. Mater. 55 (2016) 136-144.

[68] Sk. Nayab Rasool, L. Rama Moorthy, C.K. Jayasankar, Optical and luminescence properties of Dy^{3+} ions in phosphate based glasses, Solid State Sci. 22 (2013) 82-90.

[69] Sk. Mahamuda, K. Swapna, A. Srinivasa Rao , M. Jayasimhadri, T. Sasikala, K. Pavani, L. Rama Moorthy, Spectroscopic properties and luminescence behavior of Nd^{3+} doped zinc alumino bismuth borate glasses, J. Phys. Chem. Solids 74 (2013) 1308-1315.

[70] Yang Zaifa, Sun Yumei, Xu Qiguang, Sun Jiayue, Preparation and photoluminescence properties of Dy^{3+} doped $Ba_3Lu(PO_4)_3$ phosphors, J. Rare Earths 33 (2015) 1251-1255.

[71] Zhou Junhe, Yu Xue, Wang Ting, Zhou Dacheng, Qiu Jianbei, A single-phased white-emitting $Ca_2Ga_2GeO_7$:Dy^{3+} phosphor with different charge compensation ions, J. Rare Earths 35 (2017) 241-246.

[72] K. Annapoorani, P. Karthikeyan, Ch. Basavapoornima, K. Marimuthu, Investigations on the optical properties of Dy^{3+} ions doped potassium aluminium telluroborate glasses for white light applications, J. Non-Cryst. Solids 476 (2017) 128-136

[73] M. Puchalska, E. Zych, The effect of charge compensation by means of Na^+ ions on the luminescence behavior of Sm^{3+}-doped $CaAl_4O_7$ phosphor, J. Lumin. 132 (2012) 826-831.

[74] M. Venkateswarlu, Sk. Mahamuda, K. Swapna, A. Srinivasa Rao, A. Mohan Babu, Suman Shakya, D. Haranath, G. Vijaya Prakash, Luminescence spectral studies of Tm^{3+} ions

ACCEPTED MANUSCRIPT

doped Lead Tungsten Tellurite glasses for visible Red and NIR applications, J. Lumin. 175 (2016) 225-231.

[75] K. Swapna, Sk. Mahamuda, A. Srinivasa Rao, T. Sasikala, P. Packiyaraj, L. Rama Moorthy, G. Vijaya Prakash, Luminescence characterization of Eu^{3+} doped Zinc Alumino Bismuth Borate glasses for visible red emission applications, J. Lumin. 156 (2014) 80-86.

[76] C.S. Mc Camy, Correlated color temperature as an explicit function of chromaticity coordinates, Color. Res. Appl. 17 (1992) 142-144.

[77] K. Upendra Kumar, Ch. Srinivasa Rao, C. K. Jayasankar, S. Surendra Babu, J.L. Luci, Miguel A. Vallejo, Ma. Alejandrina Martinez Gamez, Optical properties of Dy^{3+}-doped P_2O_5-K_2O-MgO/MgF_2-Al_2O_3 glasses, Phys. Procedia 13 (2011) 70-73.

[78] Y.H. Zhang, D.S. Li, X. Zhao, E.Y.B. Pun, H. Lin, Photon releasing of Dy^{3+} doped fluoroborate glasses for laser illumination, J. Alloys Compd. 728 (2017) 1279-1288.

[79] T. Srihari, C.K. Jayasankar, Fluorescence properties and white light generation from Dy^{3+} doped niobium phosphate glasses, Opt. Mat. 69 (2017) 87-95.

[80] V. Himamaheswara Rao, P. Syam Prasad, M. Mohan Babu, P. Venkateswara Rao, T. Satyanarayana, Luís F. Santos, N. Veeraiah, Spectroscopic studies of Dy^{3+} ion doped tellurite glasses for solid state lasers and white LEDs, Spectrochim. Acta A 188 (2018) 516-524.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1:

Raman band positions and corresponding band assignments of OFBTDy glasses

Raman Band Position (cm^{-1})	**Raman Band Assignments**
81	Boson Peak
460	Stretching vibrations of di-borate groups and stretching vibrations of Te-O-Te linkages
780	symmetric breathing vibrations of six member rings with one or two BO_3 triangle replaced by BO_4 tetrahedral and also due to the stretching vibrations of TeO_3 units
917	Symmetric breathing vibrations of boroxol rings
984, 1028,1094, 1217	Symmetric vibrations of Tetraborate, Pyroborate, diborate and orthoborate groups respectively
1292	Stretching vibration of borate [BO_3] groups
1417, 1485, 1595	BO_2O triangle linked to BO_4 units and BO_2O triangle linked to other borate triangular units

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2:

Various physical and optical properties of Dy^{3+} ions in OFBT glasses

Physical property	**OFBTDy05**	**OFBTDy10**	**OFBTDy15**	**OFBTDy20**	**OFBTDy25**
Refractive index (n)	2.14	2.27	2.31	2.33	2.34
Density, d (g/cm^3)	2.98	3.01	3.03	3.06	3.09
Average molecular weight M (g)	84.1	85.5	86.9	88.2	89.6
Sm^{3+}-ion concentration N (x 10^{22} ions/cm^3)	1.06	2.11	3.14	4.17	5.18
Mean Atomic Volume (g/cm^3/atom)	6.72	6.76	6.82	6.86	6.9
Dielectric constant (ϵ)	4.57	5.15	5.33	5.42	5.47
Optical dielectric constant (P∂t/∂P)	3.57	4.15	4.33	4.42	4.47
Reflection losses (R)	0.131	0.150	0.156	0.159	0.160
Molar refraction R_m (cm^{-1})	15.36	16.5	16.95	17.2	17.37
Polaron radius (r_p) (Å)	1.86	1.48	1.28	1.18	1.09
Interionic distance (r_i) (Å)	5.38	4.28	3.76	3.42	3.19
Molecular electronic polarizability α (x 10^{23} cm^3)	1.21	0.65	0.48	0.34	0.27
Field Strength (x 10^{15} cm^{-2})	8.65	13.6	17.7	21.4	24.8
Optical basicity (Λ_{th})	0.525	0.528	0.531	0.534	0.536

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3:

Assignment of absorption bands, experimental (f_{exp}) (10^{-6}), calculated oscillator strengths (f_{cal}) (10^{-6}), refractive indices (n) and densities (d) of Dy^{3+} ions in OFBT glasses.

Transition	**λ**	**OFBTDy05**		**OFBTDy10**		**OFBTDy15**		**OFBTDy20**		**OFBTDy25**	
$^6H_{15/2}$ →	**(nm)**	**f_{exp}**	**f_{cal}**	**f_{exp}**	**f_{cal}**	**f_{exp}**	**f_{cal}**	**f_{exp}**	**f_{cal}**	**f_{exp}**	**f_{cal}**
$^6H_{11/2}$	1664	2.40	2.40	3.40	3.43	2.89	2.89	2.40	2.41	1.75	1.76
$^6F_{11/2}$	1264	10.0	10.1	11.0	11.1	10.0	10.6	8.42	8.13	5.42	5.09
$^6H_{7/2}$	1086	3.68	3.53	4.68	4.66	3.98	3.97	3.00	3.12	2.00	2.09
$^6F_{7/2}$	898	3.50	3.78	4.50	5.03	3.70	4.36	3.07	3.58	2.07	2.43
$^6F_{5/2}$	798	2.17	1.96	3.17	2.62	3.17	2.28	2.90	1.90	2.04	1.30
$^4F_{9/2}$	471	0.11	2.25	1.11	3.00	1.00	2.61	0.91	2.17	0.71	1.49
$^4I_{15/2}$	453	0.86	0.86	1.86	1.22	1.06	1.04	1.06	0.03	0.89	0.62
$^4G_{11/2}$	425	0.56	0.04	1.56	0.06	1.06	0.05	1.06	0.03	0.84	0.01
$^4I_{13/2}$	386	0.73	0.42	1.73	0.60	1.23	0.51	1.13	0.42	0.73	0.31
$^6P_{5/2}$	364	1.21	0.76	2.21	1.02	1.91	0.89	1.11	0.74	0.81	0.51
$^6P_{7/2}$	351	1.23	1.63	2.23	2.05	1.83	1.53	1.43	0.76	0.85	0.37
n		2.14		2.27		2.31		2.33		2.34	
d		2.9281		2.9558		2.9834		3.0111		3.0387	

ACCEPTED MANUSCRIPT

Table 4:
Comparison of Judd-Ofelt intensity parameters Ω_2, Ω_4 and Ω_6 ($10^{-20}m^2$) of Dy^{3+} ions in OFBT glasses with the other reported values.

Samples	Ω_2	Ω_4	Ω_6	Trend	Reference
OFBTDy05	8.41 ±2.02	0.65 ±0.882	2.97 ±0.38	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
OFBTDy10	12.4 ±0.07	0.73 ±0.002	3.56 ±0.005	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
OFBTDy15	9.60 ±0.62	0.51 ±0.078	3.00 ±0.097	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
OFBTDy20	7.96 ±1.77	0.23 ±0.048	2.46 ±0.772	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
OFBTDy25	6.48 ±0.04	0.10 ±0.01	1.66 ±0.012	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
DZTFB	5.28	2.54	1.97	$\Omega_2 > \Omega_4 > \Omega_6$	[45]
PKAZLFDy	7.05	1.43	1.15	$\Omega_2 > \Omega_4 > \Omega_6$	[46]
PbBAWDy	4.90	0.94	2.07	$\Omega_2 > \Omega_6 > \Omega_4$	[47]
MgTP	7.48	2.98	0.91	$\Omega_2 > \Omega_4 > \Omega_6$	[48]
D1LBE	6.45	2.02	3.87	$\Omega_2 > \Omega_6 > \Omega_4$	[49]
DyZB10P	3.95	0.91	1.95	$\Omega_2 > \Omega_6 > \Omega_4$	[50]
PbFPDy	7.12	1.59	2.20	$\Omega_2 > \Omega_6 > \Omega_4$	[51]
NbFSDy	10.05	1.37	2.16	$\Omega_2 > \Omega_6 > \Omega_4$	[52]
Oxyfluoroborate	2.68	2.56	0.89	$\Omega_2 > \Omega_4 > \Omega_6$	[53]
BZBDy	4.03	1.14	1.65	$\Omega_2 > \Omega_6 > \Omega_4$	[54]

ACCEPTED MANUSCRIPT

Table 5:

Transition probabilities (A_R) (s^{-1}), total transition probability (A_T) (s^{-1}), radiative branching ratio (β_R) and radiative lifetimes (τ_R)(μs) for the observed emission transitions of Dy^{3+} ions in OFBT glasses

Name	Transition	A_R	A_T	β_R	τ_R
OFBTDy05	$^4F_{9/2} \rightarrow {}^6H_{15/2}$	661.88 ±110	3835.12 ±198	0.1726 ±0.04	260 ±15
	$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2500.97 ±275		0.6521 ±0.04	
OFBTDy10	$^4F_{9/2} \rightarrow {}^6H_{15/2}$	990.90 ±2	6441.15 ±8	0.1838 ±0.03	155 ±0
	$^4F_{9/2} \rightarrow {}^6H_{13/2}$	4382.01 ±5		0.6803 ±0.004	
OFBTDy15	$^4F_{9/2} \rightarrow {}^6H_{15/2}$	889.74 ±22	5519.30 ±166	0.1612 ±0.009	181 ±5
	$^4F_{9/2} \rightarrow {}^6H_{13/2}$	3685.41 ±154		0.6677 ±0.008	
OFBTDy20	$^4F_{9/2} \rightarrow {}^6H_{15/2}$	744.59 ±228	4718.36 ±1127	0.1578 ±0.001	211 ±67
	$^4F_{9/2} \rightarrow {}^6H_{13/2}$	3133.51 ±764		0.6641 ±0.004	
OFBTDy25	$^4F_{9/2} \rightarrow {}^6H_{15/2}$	509.71 ±1.35	3706.53 ±15	0.1375 ±0.001	269 ±1
	$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2473.99 ±13		0.6675 ±0.001	

ACCEPTED MANUSCRIPT

Table 6:

Emission peak wavelength (λ_P)(nm), effective band widths ($\Delta\lambda_P$)(nm), experimental branching ratios (β_{exp}), stimulated emission cross-sections (σ_{se}) ($x10^{-22}$) (cm^2), gain band width (σ_{se} x $\Delta\lambda_p$) (10^{-27}) (cm^3) and optical gain parameters (σ_{se} x τ_R) (10^{-25}) (cm^2 s) for the emission transitions of Dy^{3+} ions in OFBT glasses.

Spectral parameters	**OFBTDy05**	**OFBTDy10**	**OFBTDy15**	**OFBTDy20**	**OFBTDy25**
$^4F_{9/2}\rightarrow{}^6H_{15/2}$					
λp	485	485	485	485	485
$\Delta\lambda p$	17.3	17.65	17.61	16.09	16.86
β_{exp}	0.1632	0.1742	0.1532	0.1446	0.1420
σ_{se}	6.13 ±1.02	8.16 ±0.016	7.08 ±0.18	5.82 ±1.78	3.95 ±0.016
σ_{se} x $\Delta\lambda p$	1.06	1.41	1.22	1.01	0.68
σ_{se} x τ_R	1.59	1.26	1.28	1.23	1.06
$^4F_{9/2}\rightarrow{}^6H_{13/2}$					
λp	576	576	576	576	576
$\Delta\lambda p$	16.23	16.28	15.62	16.51	16.51
β_{exp}	0.5821	0.6578	0.6234	0.6321	0.6241
σ_{se}	49.1 ±5.40	76.5 ±0.09	62.2 ±2.60	51.9 ±12.6	40.7 ±0.22
σ_{se} x $\Delta\lambda p$	7.98	12.4	10.1	8.43	6.60
σ_{se} x τ_R	12.8	11.9	11.3	11.0	10.9

ACCEPTED MANUSCRIPT

Table 7:

Radiative lifetimes (τ_R) (μs), experimental lifetimes (τ_{exp}) (μs), quantum efficiency (η) and non-radiative relaxation rates (W_{NR}) $(\mu s)^{-1}$ for Dy^{3+} ions in OFBT glasses.

Sample	τ_R	τ_{exp}	η (%)	W_{NR}	References
OFBTDy05	260 ±15	175 ±4.9	67.32 ±3.68	1868	Present work
OFBTDy10	155 ±0	142 ±4.0	91.61 ±0.29	591	Present work
OFBTDy15	181 ±5	123 ±1.5	67.95 ±1.82	2605	Present work
OFBTDy20	211 ±67	113 ±2.8	53.55 ±2.90	4110	Present work
OFBTDy25	269 ±1	109 ±4.9	40.50 ±0.12	5457	Present work

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 8:

CIE chromaticity co-ordinates, Y/B intensity ratios and color correlated temperature (CCT) (K) of Dy^{3+} ions in OFBT glasses

Name of the glass	Color Co-ordinates		Y/B Intensity Ratio	CCT (K)
	X-Co-ordinates	Y-Co-ordinates		
OFBTDy05	0.3096	0.3262	0.97	6700
OFBTDy10	0.3364	0.3402	1.11	5328
OFBTDy15	0.3032	0.3604	1.02	6743
OFBTDy20	0.3169	0.3706	0.90	6101
OFBTDy25	0.3381	0.4019	0.86	5330

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 9:

Comparison of branching ratio (β_R), stimulated emission cross-section (σ_{se}) (10^{-22}), quantum efficiency (η) and Y/B intensity ratio with other reported values.

Sample	β_R	σ_{se}	η (%)	Y/B	References
OFBTDy10	0.68 ±0.0004	76.5 ±0.09	91 ±0.29	1.11	Present work
YCaSBDy0.5	0.58	9.53	-	1.27	[12]
PKMAFDy10	0.59	2.49	-	1.55	[77]
NCZLB0.05	0.63	-	75	-	[78]
PKANBDy	0.55	6.40	92	1.94	[79]
TSWD10	0.73	58.3	69	1.49	[80]

ACCEPTED MANUSCRIPT

Figures

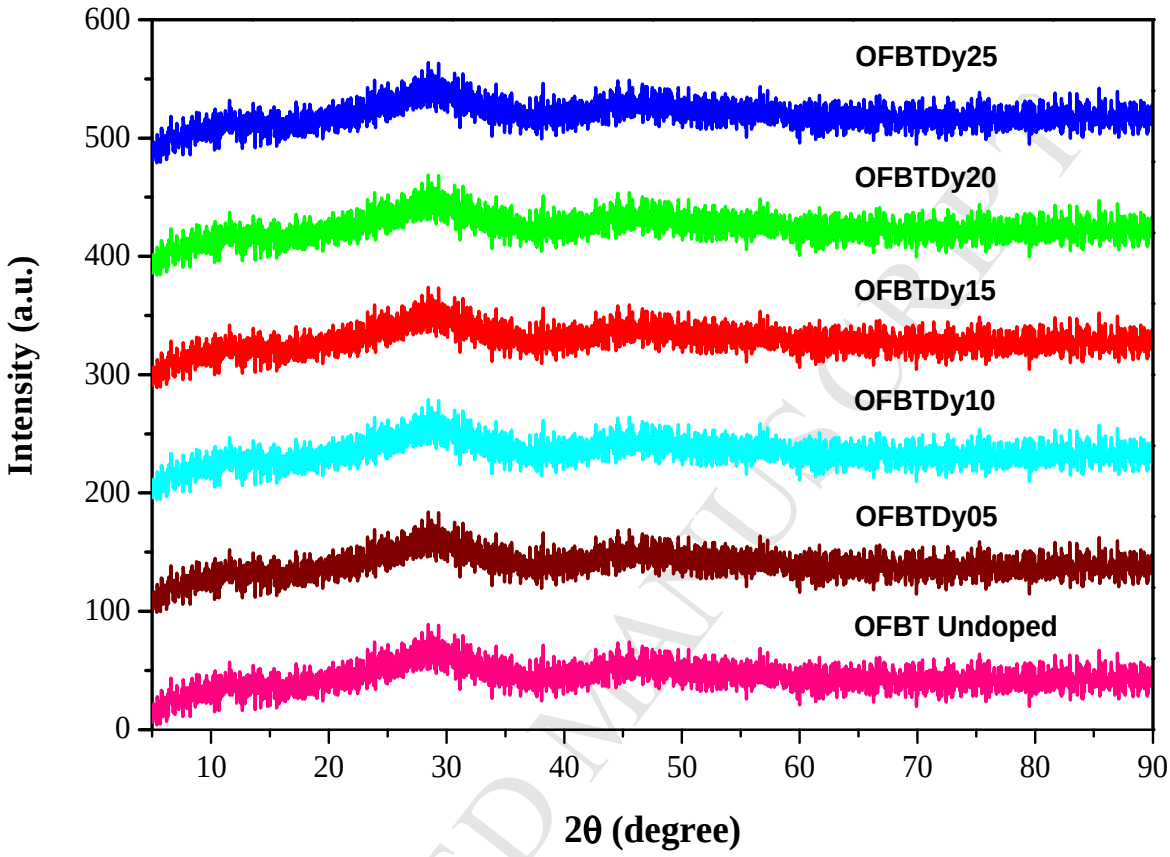


Fig. 1: XRD spectra of an un-doped and Dy^{3+} ions doped OFBT glasses

ACCEPTED MANUSCRIPT

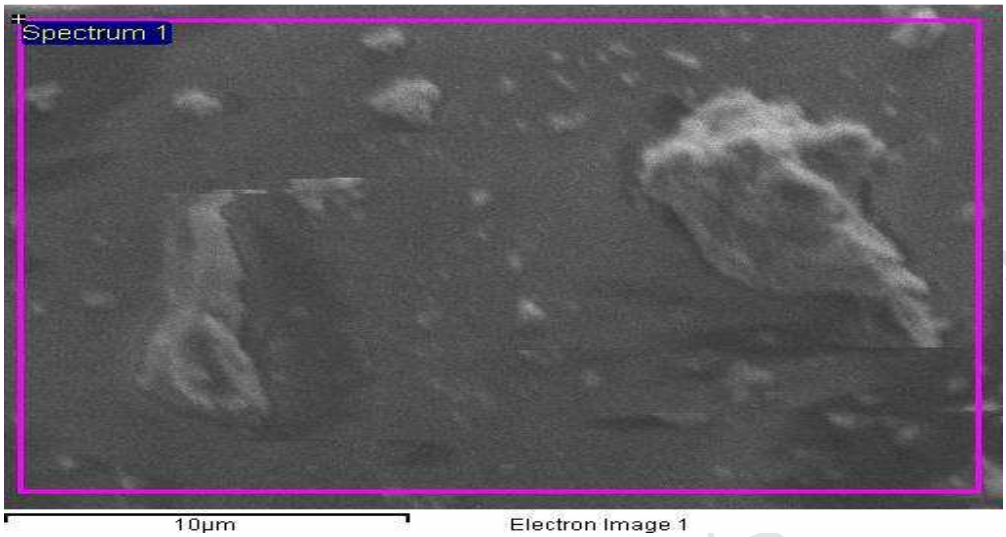


Fig. 2: SEM image of un-doped OFBT glass

ACCEPTED MANUSCRIPT

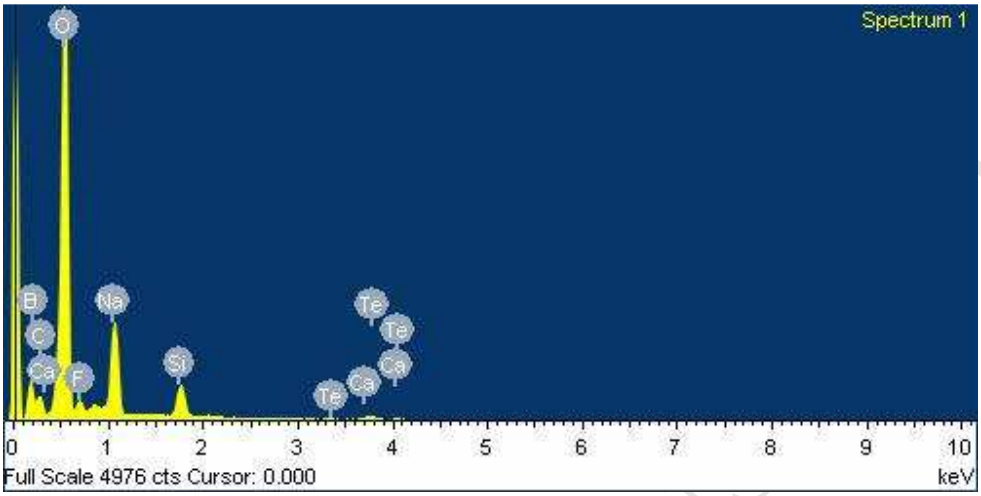


Fig. 3: EDAX spectrum of un-doped OFBT glass

ACCEPTED MANUSCRIPT

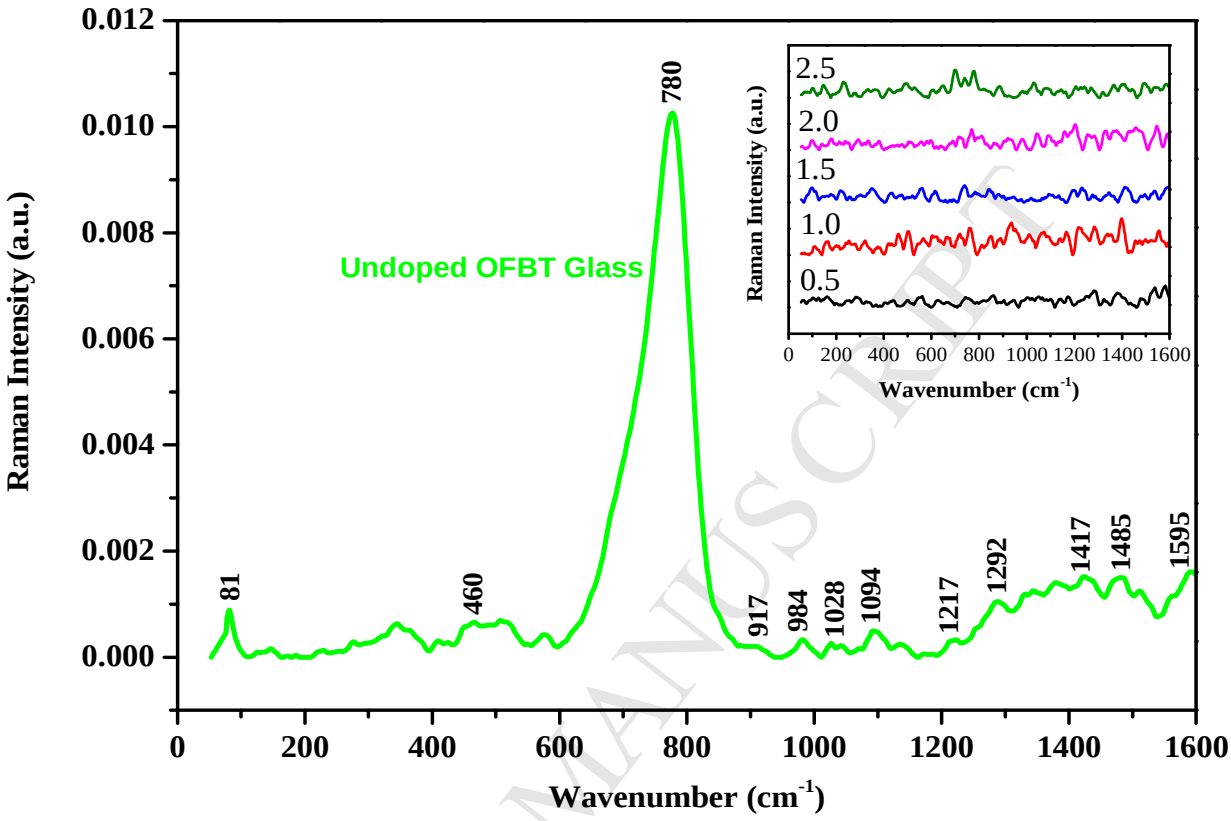


Fig. 4: Raman spectrum of an un-doped OFBT glass (Inset figure shows the Raman spectra of Dy^{3+} ions doped glasses)

ACCEPTED MANUSCRIPT

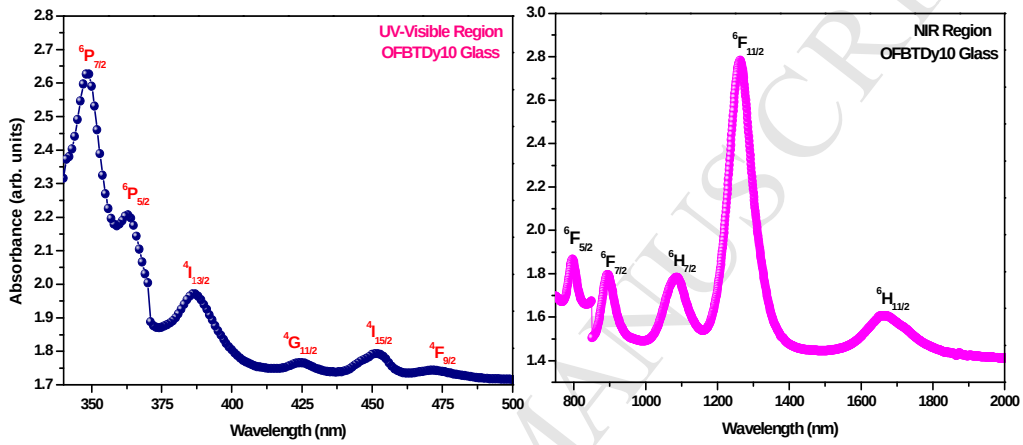


Fig. 5: Absorption spectrum of OFBTDy10 glass in UV-visible and NIR regions

ACCEPTED MANUSCRIPT

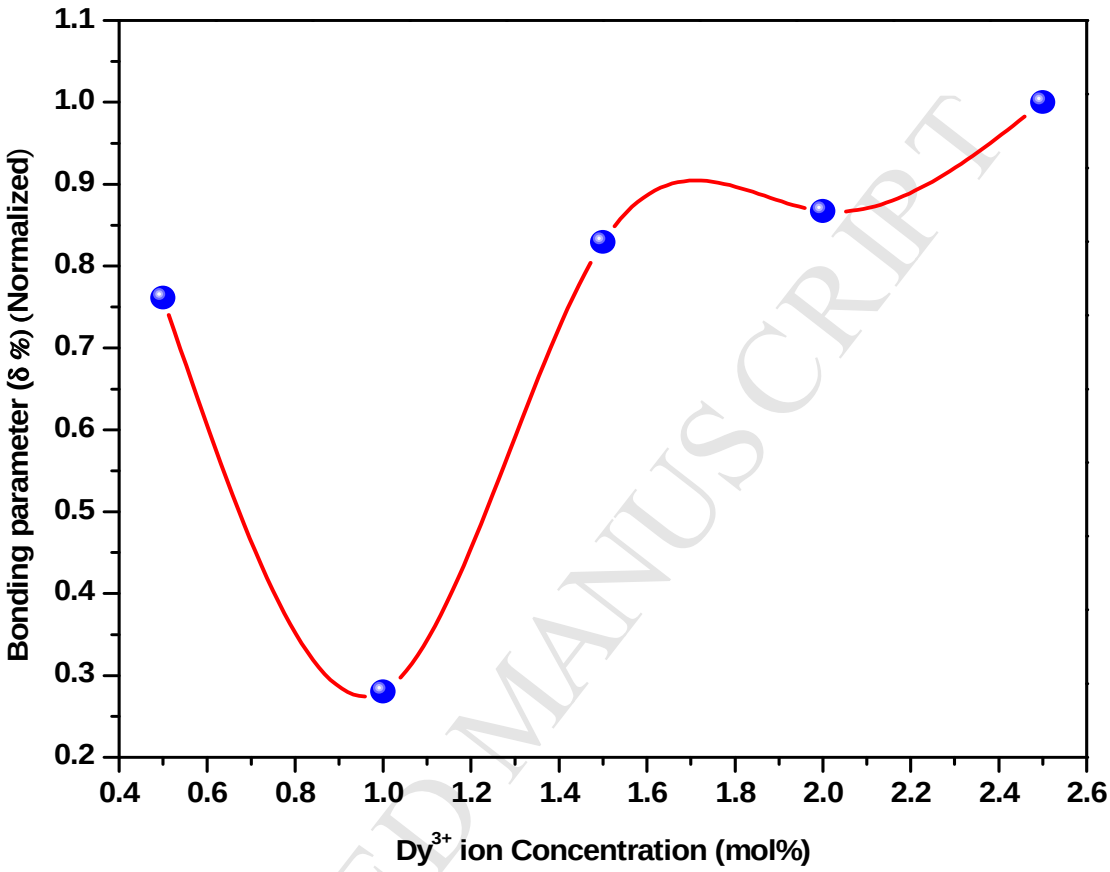


Fig. 6: Variation of bonding parameter with Dy^{3+} ion concentration of OFBT glasses

ACCEPTED MANUSCRIPT

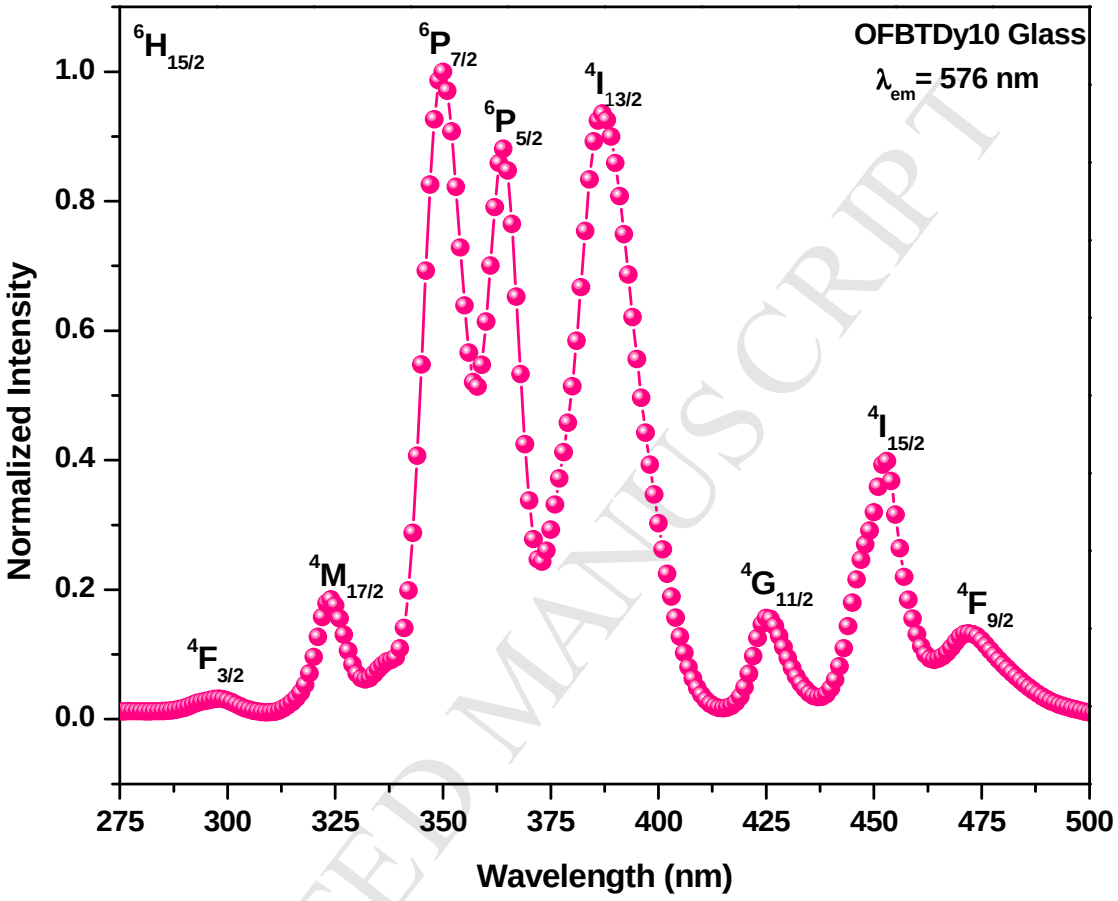


Fig. 7: Excitation spectrum of OFBTDy10 glass

ACCEPTED MANUSCRIPT

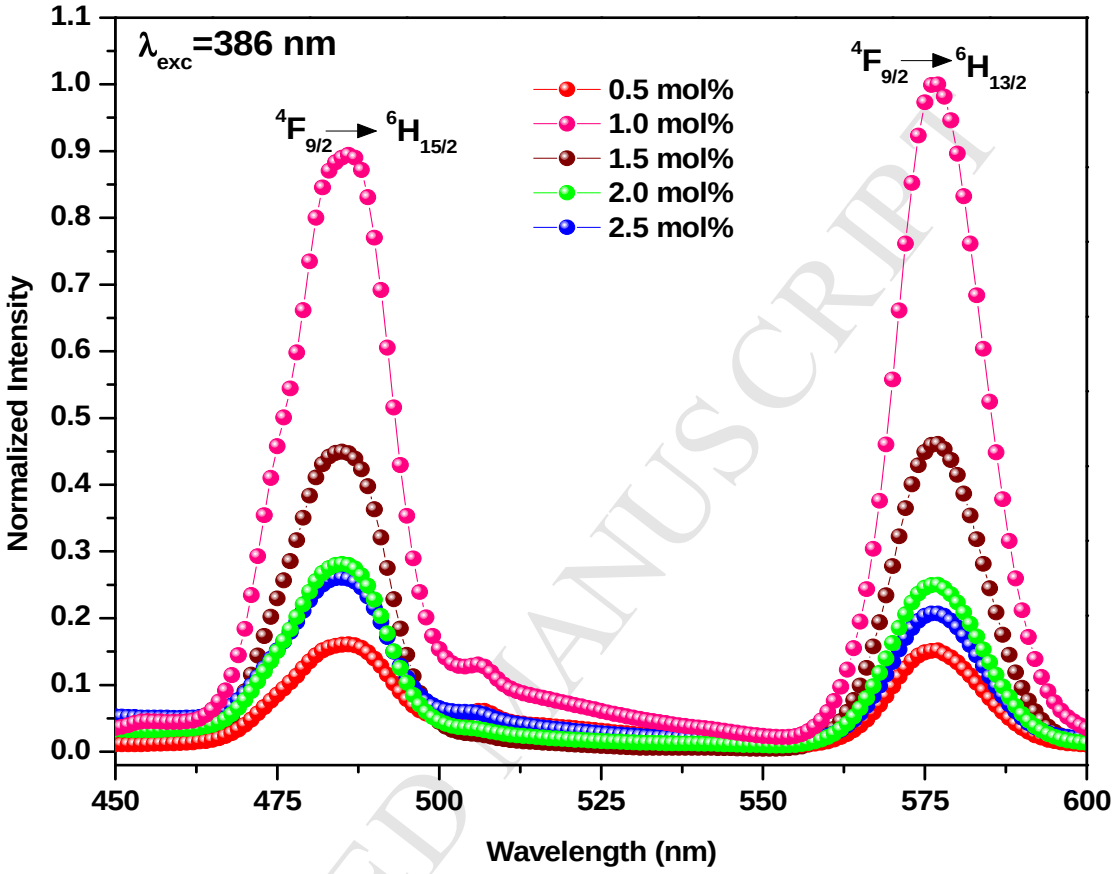


Fig. 8: PL spectra of different concentrations of Dy^{3+} ions in OFBT glasses

ACCEPTED MANUSCRIPT

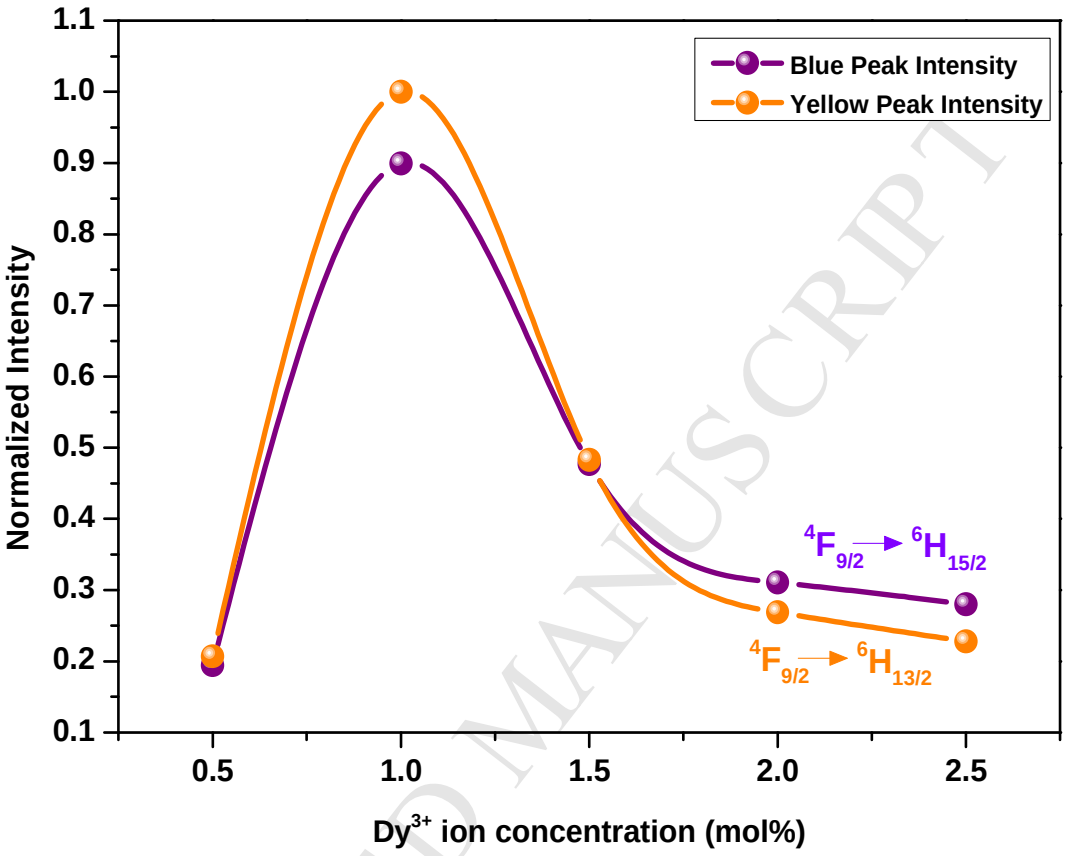


Fig. 9: Variation of yellow and blue peaks intensities with Dy^{3+} ions concentration in OFBT glasses

ACCEPTED MANUSCRIPT

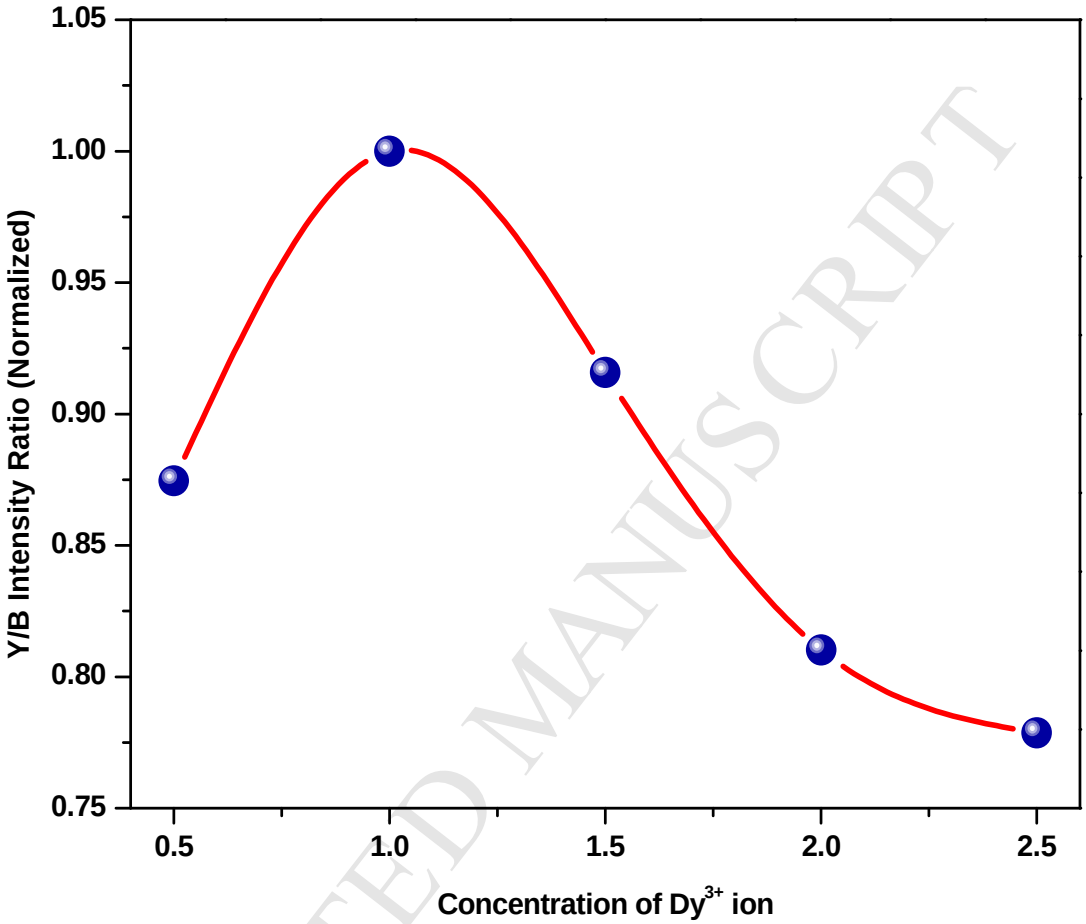


Fig. 10: Variation of Y/B intensity ratio with the concentration of Dy^{3+} ions in OFBT glasses

ACCEPTED MANUSCRIPT

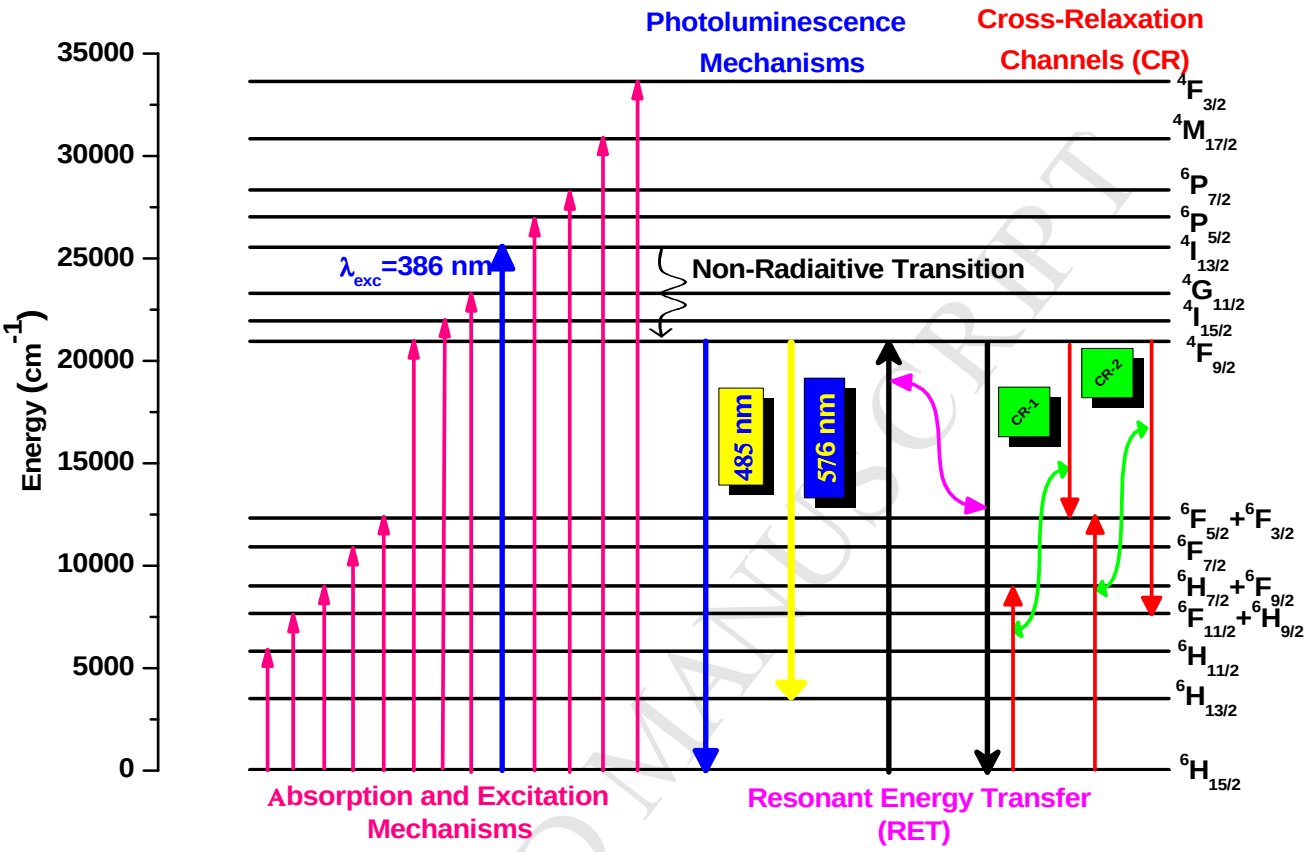


Fig. 11: Energy level diagram of 1 mol% of Dy^{3+} ions doped OFBT glass with various mechanisms.

ACCEPTED MANUSCRIPT

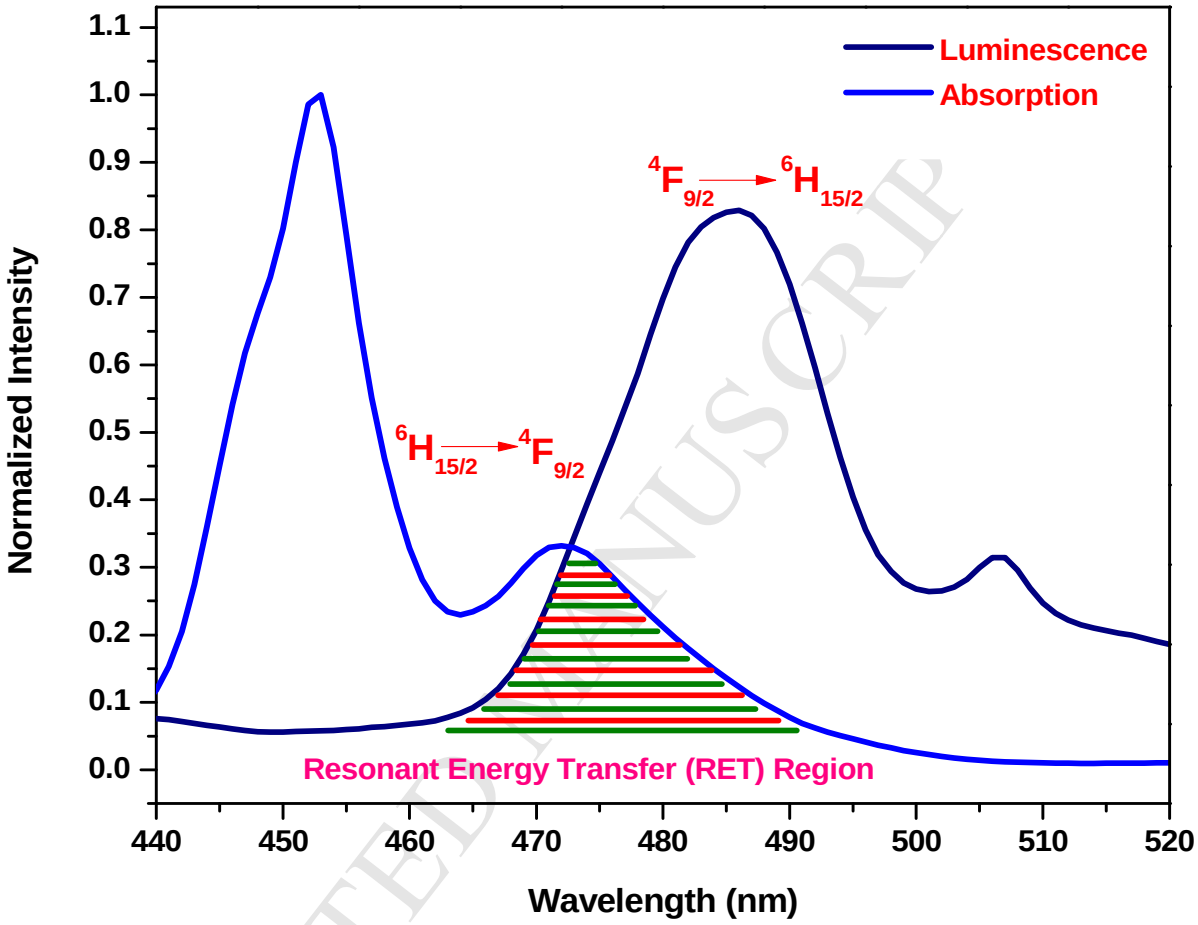


Fig. 12: Resonance energy transfer (RET) region of the Dy^{3+}ions in OFBT glasses

ACCEPTED MANUSCRIPT

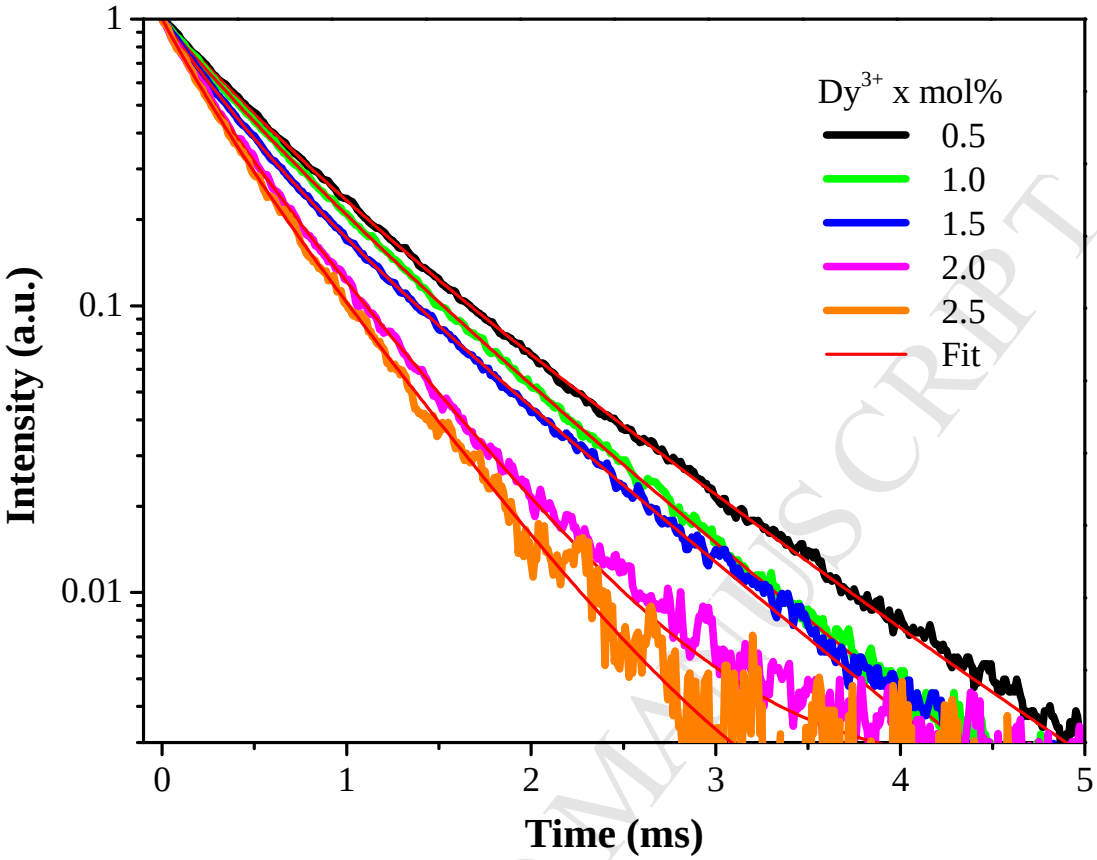


Fig. 13: PL decay curves of Dy^{3+} ions in OFBT glasses with fitting.

ACCEPTED MANUSCRIPT

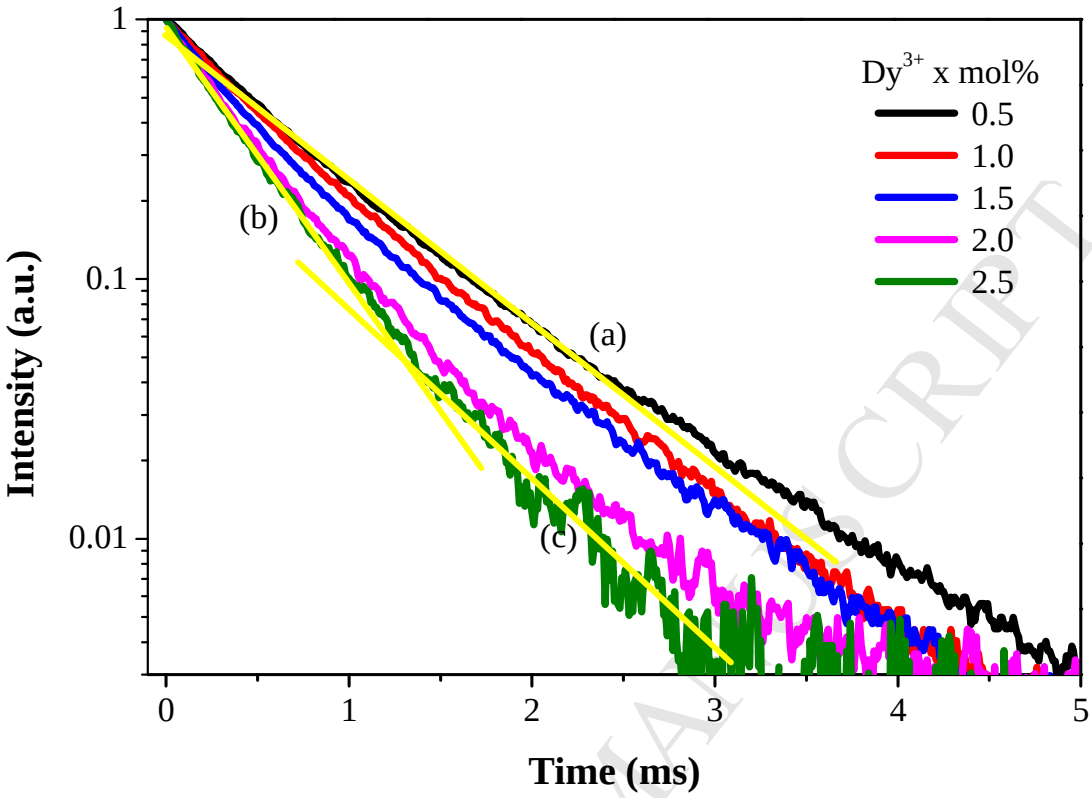


Fig. 14: PL decay curves of Dy^{3+} ions doped OFBT glasses which shows the conversion of exponential to non-exponential behavior with increase in Dy^{3+} ion concentration

ACCEPTED MANUSCRIPT

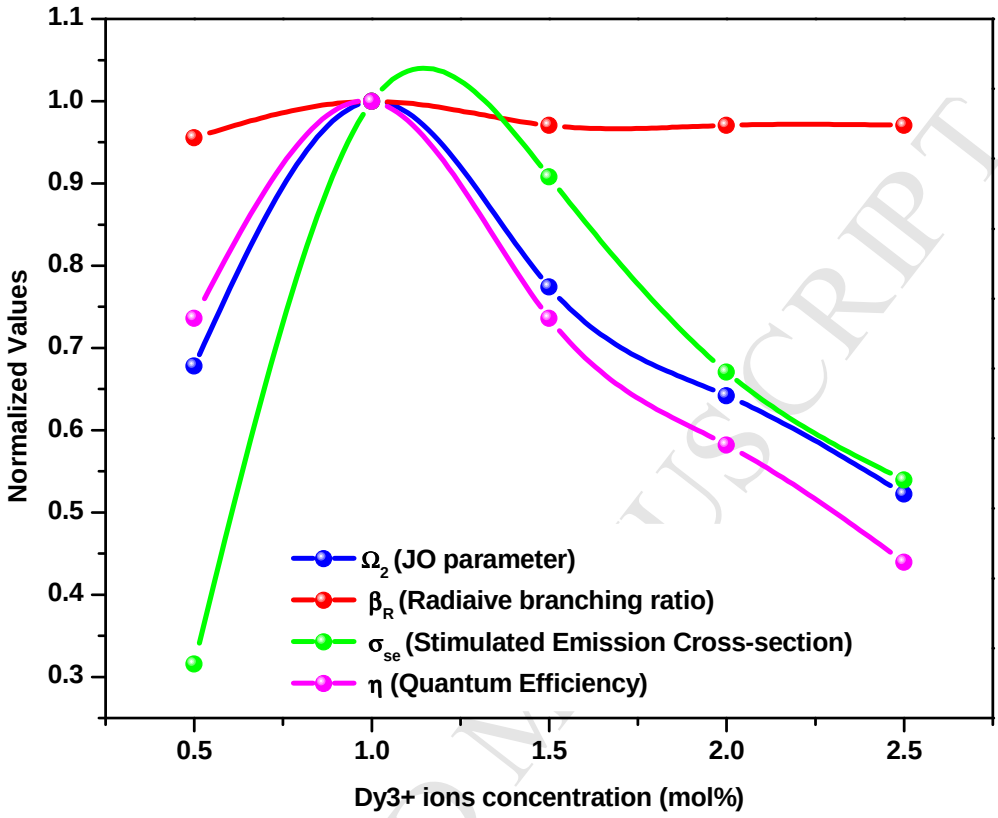


Fig. 15: Variation of J-O parameter (Ω_2), radiative branching ratio (β_R), stimulated emission cross-section (σ_{se}) and quantum efficiency (η) with the Dy^{3+} ion concentration in OFBT glasses.

ACCEPTED MANUSCRIPT

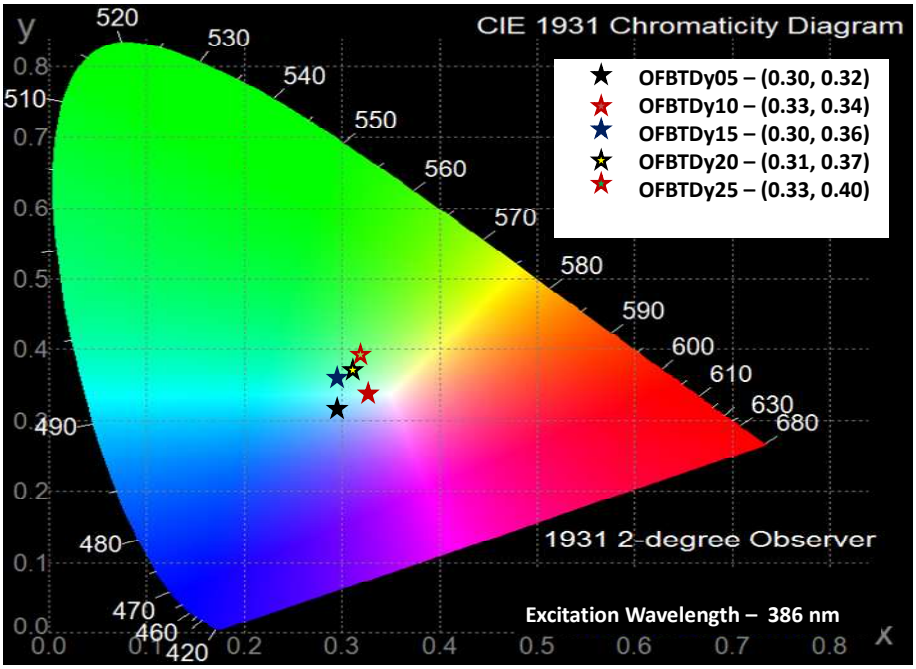


Fig. 16: CIE Color Chromaticity diagram of Dy^{3+} ions doped OFBT glasses.

ACCEPTED MANUSCRIPT

Highlights

- Dy^{3+} ions doped OFBT glasses were prepared by melt quenching technique
- An intense emission is observed in the yellow region under 386 nm excitation.
- Y/B intensity ratio is measured to analyze the structure of the luminescent glass.
- OFBTDy10 glass is aptly suitable for solid state lighting applications.

ACCEPTED MANUSCRIPT

Differ Equ Dyn Syst
https://doi.org/10.1007/s12591-018-0424-8

ORIGINAL RESEARCH

Dynamical Model of Epidemic Along with Time Delay; Holling Type II Incidence Rate and Monod–Haldane Type Treatment Rate

Abhishek Kumar[1] · **Nilam**[1]

© Foundation for Scientific Research and Technological Innovation 2018

Abstract The present study aims to control the infectious diseases and epidemics in the human population. Therefore, in the present article, we have proposed a delayed SIR epidemic model along with Holling type II incidence rate and treatment rate as Monod–Haldane type. Model stability has been established in the three regions of the basic reproduction number R_0 i.e. R_0 equals to one, greater than one and less than one. The model is locally asymptotically stable for disease-free equilibrium Q when the basic reproduction number R_0 is less than one ($R_0 < 1$) and unstable when $R_0 > 1$ for time lag $\tau \geq 0$. We investigated the stability of the model for disease-free equilibrium at R_0 equals to one using central manifold theory. Using center manifold theory, we proved that at $R_0 = 1$, disease-free equilibrium changes its stability from stable to unstable. We also investigated the stability for endemic equilibrium Q^* for time lag $\tau \geq 0$. Further, numerical simulations are presented to exemplify the analytical studies.

Keywords Epidemic · SIR model · Delay differential equation · Monod–Haldane type treatment rate · Holling type II incidence rate · Stability · Center manifold theory

Mathematics Subject Classification 34D20 · 92B05 · 37M05

Introduction

In mathematical epidemiology literature many authors [1–3, 10, 13, 14] have suggested the various models for the disease transmission like susceptible–infected–recovered (SIR) model, susceptible–infected–recovered–susceptible (SIRS) and many more. In the population

✉ Nilam
rathi.nilam@gmail.com

Abhishek Kumar
abhishek.dtu14@gmail.com

[1] Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India

Published online: 10 May 2018

dynamics, both transmission and treatment always play an important role. Transmission of infection is the process in which susceptible are getting infected via infected population through various channels [11]. Several authors have suggested that the disease transmission process may have a nonlinear incidence rate. In many epidemic models [1, 9], the bilinear incidence rate βSI and the standard incidence rate $\frac{\beta SI}{N}$ are frequently used. In literature, the bilinear incidence rate is based on the law of mass action. The standard incidence rate may be a good approximation and it is practically feasible if the number of available partners is large enough and everybody could not make more contacts. In fact, the infection probability per contact is likely influenced by the number of infective individuals, because more infective individuals can increase the infection risk [13]. Capasso and Serio [5] introduced a saturated incidence rate g(I)S into epidemic models, where g(I) tends to a saturation level when I gets large, i.e.

$$g(I) = \frac{\beta I}{1 + \alpha I},$$

where βI measures the infecting force of the disease and $\frac{1}{1+\alpha I}$ measures the inhibition effect of the behavioural change of the susceptible individuals when their number increase or from the crowding effect of the infective individuals. The incidence rate g (I) S with $g(I) = \frac{\beta I}{1+\alpha I}$ is also known as Holling type II incidence rate. This incidence rate seems more reasonable than the bilinear incidence rate βIS, because it includes the behavioural change and crowding effect of the infective individuals and prevents the unboundedness of the contact rate by choosing suitable parameters [13].

In the field of epidemiology, treatment, vaccination, and many more play an important role in controlling the disease spread. Recently, many researchers [3, 4, 11] have focused on nonlinear type treatment rate. Different type of treatment rates like Holling type II, Holling type III [3, 4, 11] have been implemented by the authors in their model to study the dynamics of infectious diseases. In 1968, Andrews [8] had suggested a function having the form

$$G(I) = \frac{mI}{a + bI + I^2},$$

called the Monod–Haldane function. Sokol and Howell [18] had proposed a simplified Monod–Haldane function having the form

$$G(I) = \frac{mI}{a + I^2}.$$

In 2009, Baek et al. [7] used this simplified Monod–Haldane function to describe the dynamical relation between prey and predator.

Considering these facts we have incorporated treatment rate as Monod–Haldane function type in our delayed SIR model with Holling type II incidence rate.

In the present article, we have analyzed the effect of delay on the SIR model along with nonlinear Holling Type II incidence rate and Monod–Haldane type treatment rate for the better understanding of disease dynamics. Furthermore, we evaluate the basic reproduction number R_0, analyzed the dynamical behavior of the model and also discussed the stability of the model. The basic reproduction number R_0 is defined as "the expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual" [12]. The stability analysis of the model has been done by Descartes's Rule [19] along with Routh–Hurwitz criterion.

The rest of the paper is organized as follows: after the abstract and introduction, the mathematical model is formulated in Sect. 2. Basic properties of solutions are discussed in

Sect. 3. Equilibrium and stability analysis of the model is in Sect. 4. Numerical simulation and discussion are in Sect. 5. Finally, the conclusion is summarized in Sect. 6.

Mathematical Model

We assume that the total population N is divided into three compartments: susceptible individuals compartment S(t), infected individuals compartment I(t) and recovered individuals compartment R(t). Susceptible individuals are those who can get a disease under appropriate conditions. Infected individuals are the one who has got a disease and can spread the disease to susceptible individuals via contacts. As time passes, infected individuals lose infectivity and move to recover compartment by auto recovery due to autoimmune response of the body or by treatment. We considered simplified Monod–Haldane type treatment rate for the recovery of the infected population. The progression of an epidemic in a different compartment shown by block diagram below (Fig. 1).

The rate of change of the population in each compartment is given by the following nonlinear system of the delay differential equations:

$$\frac{dS}{dt} = A - \mu S - \frac{\beta\, S(t-\tau) I(t-\tau)}{1+\alpha\, I(t-\tau)}, \tag{1}$$

$$\frac{dI}{dt} = \frac{\beta\, S(t-\tau)\, I(t-\tau)}{1+\alpha\, I(t-\tau)} - (\mu + d + \delta)\, I - \frac{aI}{I^2+b}, \tag{2}$$

$$\frac{dR}{dt} = \frac{aI}{I^2+b} + \delta\, I - \mu\, R. \tag{3}$$

where $\tau > 0$ is a fixed time during which the infectious agents develop in the vector and it is only after that time that the infected vector can infect a susceptible individual.

Let $C = C\left([-\tau, 0], \mathbb{R}^3\right)$ be the Banach space of continuous functions mapping the interval $[-\tau, 0]$ to $\mathbb{R}^3$ with the topology of uniform convergence. By the fundamental theory of functional differential equations [10], it can be shown that there exists a unique solution $(S(t), I(t), R(t))$ of model (Eqs. 1–3) with initial data $(S_0, I_0, R_0) \in C$. For ecological reasons, we assume that the initial conditions of model (Eqs. 1–3) satisfy:

$$S_0(\varphi) \geq 0,\ I_0(\varphi) \geq 0,\ R_0(\varphi) \geq 0,\ \varphi \in [-\tau, 0]. \tag{4}$$

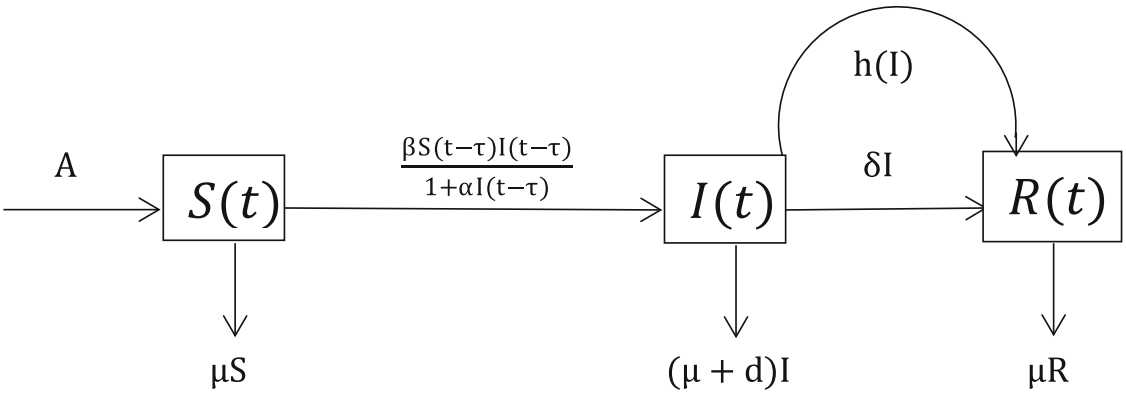


Fig. 1 Progression of infection from susceptible (**S**) individuals through infected (**I**) and recovered (**R**) compartments for the model

Table 1 Parameters and description

Parameter	Description
A	Recruitment rate per day
α	Inhibition rate due to infected
β	Transmission rate
μ	Natural mortality rate
d	Disease induced mortality rate
δ	Recovery rate
a	Treatment rate
b	Limitation rate in treatment availability

The term, $\mathrm{h}(\mathrm{I}) = \frac{\mathrm{aI}}{\mathrm{b}+\mathrm{I}^2}$ in the system (1)–(3) represents the simplified Monod–Haldane type treatment term, where a is treatment rate and b is limitation rate in treatment availability. In Monod Haldane, the removal/treatment rate initially grows with the growth of infectives and reaches the maximum and then starts decaying. Such a situation may arise due to limitation in the availability of treatment for a large number of infected individuals. When supplies of treatment (medicine, immunization, etc.) are depleted, then in spite of the high number of infectives the available treatments become very low. This case may arise when there is the re-emergence and spread of disease in place of limited treatment facilities [3]. The term $\frac{\beta\, \mathrm{S}(\mathrm{t}-\tau)\mathrm{I}(\mathrm{t}-\tau)}{1+\alpha\, \mathrm{I}(\mathrm{t}-\tau)}$ in the system (1)–(3) represent the Holling type II incidence rate. The incidence rate $\frac{\beta\, \mathrm{S}(\mathrm{t}-\tau)\mathrm{I}(\mathrm{t}-\tau)}{1+\alpha\, \mathrm{I}(\mathrm{t}-\tau)}$ represents the rate at time $(\mathrm{t}-\tau)$ at which susceptible individuals leave the susceptible class and enter in the infectious class at time t (Table 1).

Basic Properties of the Model

From the system (1)–(3) we can infer that S and I are free from the effect of R. Thus it is enough to consider the following reduced system for study:

$$\frac{\mathrm{dS}}{\mathrm{dt}} = \mathrm{A} - \mu\, \mathrm{S} - \frac{\beta\, \mathrm{S}(\mathrm{t}-\tau)\mathrm{I}\,(\mathrm{t}-\tau)}{1+\alpha\, \mathrm{I}\,(\mathrm{t}-\tau)}, \tag{5}$$

$$\frac{\mathrm{dI}}{\mathrm{dt}} = \frac{\beta\, \mathrm{S}\,(\mathrm{t}-\tau)\,\mathrm{I}\,(\mathrm{t}-\tau)}{1+\alpha\, \mathrm{I}\,(\mathrm{t}-\tau)} - (\mu+\mathrm{d}+\delta)\,\mathrm{I} - \frac{\mathrm{aI}}{\mathrm{I}^2+\mathrm{b}}, \tag{6}$$

with initial conditions

$$S_0\,(\varphi) \geq 0,\, I_0\,(\varphi) \geq 0,\, \varphi \in [-\tau, 0]\,. \tag{7}$$

The system (5)–(6) monitors the population. From [10] it can be shown that all state variables of the model (5)–(6) are nonnegative i.e. $(\mathrm{S}, \mathrm{I}) \in \mathbb{R}^2_+$. It is assumed that the parameters $\mathrm{A}, \mu, \beta, \mathrm{d}, \delta, \mathrm{a}, \alpha, \mathrm{b} > 0$.

Theorem 1 *Model (5)–(6) has a nonnegative solution with the initial value (7).*

Proof First we show that $S(t)$ is nonnegative for all $t \geq 0$. On the contrary, it is assumed that there exist $t_1 > 0$ be the first time such that $S\,(t_1) = 0$, then by the Eq. (5) we have $S^{'}\,(t_1) = A > 0$, and hence $S\,(t) < 0$ for $t \in (t_1 - \theta, t_1)$, where $\theta > 0$ is sufficiently small. This contradicts $S\,(t) > 0$ for $t \in [0, t_1)$. It follows that $S\,(t) > 0$ for $t > 0$. Now, we prove

that positivity of solution $I(t)$. By using the variation of constant formula and the step by step integration method, integrating the Eq. (6) of the model from 0 to t for $0 < t \leq \tau$, we obtain

$$I\,(t) = I(0)e^{-(\mu+d+\delta)t} \times e^{\int_0^t f(S(\varepsilon-\tau),\mathrm{I}(\varepsilon-\tau),I(\varepsilon))d\varepsilon}.$$

Here, $f\,(S\,(\varepsilon-\tau)\,,\mathrm{I}\,(\varepsilon-\tau)\,,I\,(\varepsilon)) = \left(\frac{\beta\,S(\varepsilon-\tau)\mathrm{I}(\varepsilon-\tau)}{(1+\alpha\,\mathrm{I}(\varepsilon-\tau))I(\varepsilon)} - \frac{a}{\mathrm{I}^2(\varepsilon)+b}\right)$

It is easy to see that $I\,(t) > 0$ for all $0 \leq t \leq \tau$. Then integrating the Eq. (6) of the model from τ to t for $\tau < t \leq 2\tau$ gives

$$I\,(t) = I(0)e^{-(\mu+d+\delta)t} \times e^{\int_\tau^t f(S(\varepsilon-\tau),\mathrm{I}(\varepsilon-\tau),I(\varepsilon))d\varepsilon}.$$

Note that $I\,(t) > 0$ for all $\tau \leq t \leq 2\tau$ and this process can easily carry on. It implies that for all $t > 0$, we have $I\,(t) > 0$.

Next we prove that the solution of model ultimately uniformly bounded for all $t \geq 0$.

It is convenient to define $\mathrm{N}_0 = \frac{\mathrm{A}}{\mu}$, $\mathrm{S}_0 = \frac{\mathrm{A}}{\mu}$. The rate of change of the total population, obtained by adding all Eqs. (1)–(3), is given by:

$$\frac{\mathrm{dN}}{\mathrm{dt}} = \mathrm{A} - \mu\,\mathrm{N} - \mathrm{dI} \leq \mathrm{A} - \mu\,\mathrm{N}. \tag{8}$$

With $\mathrm{N} = \mathrm{S} + \mathrm{I} + \mathrm{R}$.

Since as $\mathrm{t} \to \infty$, disease will disappear, therefore $\limsup_{\mathrm{t}\to\infty} \mathrm{I}\,(\mathrm{t}) = 0$,

Equation (8) $\Rightarrow \frac{\mathrm{dN}}{\mathrm{dt}} = \mathrm{A} - \mu\,\mathrm{N}$

$$\Rightarrow \mathrm{N}\,(\mathrm{t}) = \frac{\mathrm{A}}{\mu} + \left(\mathrm{N}(0) - \frac{\mathrm{A}}{\mu}\right)\mathrm{e}^{-\mu\,\mathrm{t}},$$

$$\Rightarrow \lim_{\mathrm{t}\to\infty} \mathrm{N}\,(\mathrm{t}) = \frac{\mathrm{A}}{\mu} = \mathrm{N}_0.$$

This follows that

$0 < \limsup_{\mathrm{t}\to\infty} \mathrm{N}\,(\mathrm{t}) \leq \mathrm{N}_0$ if and only if $\limsup_{\mathrm{t}\to\infty} \mathrm{I}\,(\mathrm{t}) = 0$

From the Eq. (5), it follows that

$$0 < \limsup_{\mathrm{t}\to\infty} \mathrm{S}\,(\mathrm{t}) \leq \mathrm{S}_0. \tag{9}$$

From the Eq. (8) it follows that if $\mathrm{N} > \mathrm{N}_0$ then $\frac{\mathrm{dN}}{\mathrm{dt}} < 0$. This establishes the following lemma.

Lemma 1 *The set*

$$D = \{(S, I) \in \mathbb{R}^2_+ : S + I \leq N_0, S \leq S_0\},$$

is a positively invariant and attracting region for the disease transmission model given by the system (5)–(6) with initial conditions (7) in $\mathbb{R}^2_+$

Consequently, in the absence of disease ($\mathrm{I} = 0$), the total population N approaches the carrying capacity N_0 asymptotically, and in the presence of disease, the total population is less than or equals to N_0.

Thus, every solution of the system (5)–(6) with initial conditions in $\mathbb{R}^2_+$ tends towards D as $\mathrm{t} \to \infty$, and every solution with an initial condition in D remains there for $\mathrm{t} > 0$. Furthermore, in D, the usual existence, uniqueness and continuation results hold for the system so that the model (5)–(6) is well posed mathematically and epidemiologically.

Equilibriums and Stability Analysis

Equilibria of the system (5)–(6) are obtained by setting the right-hand sides of the equations of the system to zero. Let $Q^* = (S^*, I^*)$ be any equilibrium point of the model. Naturally a stable equilibrium point Q^* is the one for which all the nearby trajectories approach it asymptotically as $t \to \infty$. Hence, the equilibrium solutions of a system with time delays are the same as those of the corresponding system with zero delay [17].

The Disease-Free Equilibrium

The system (5)–(6) has a unique disease-free equilibrium of the form $Q = (\frac{A}{\mu}, 0)$, obtained by setting the right-hand sides of the equation of system to zero.

Then the characteristic equation at Q of the system (5)–(6) is given by

$$(\mu + \lambda)\left(\frac{\beta A}{\mu}e^{-\lambda\tau} - \mu - d - \delta - \frac{a}{b} - \lambda\right) = 0. \tag{10}$$

One of the roots of Eq. (10) is given by $\lambda_1 = -\mu$ and other can be obtained from

$$\left(\frac{\beta A}{\mu}e^{-\lambda\tau} - \mu - d - \delta - \frac{a}{b} - \lambda\right) = 0,$$

where,

The term $\frac{\beta A}{\mu(\mu+d+\delta+\frac{a}{b})}e^{-\lambda\tau}$ at $\tau = 0$, is known as the basic Reproduction number, denoted by R_0. Therefore, we define the basic reproduction number R_0 of our model by

$$R_0 = \frac{\beta A}{\mu\left(\mu + d + \delta + \frac{a}{b}\right)},$$

Where $\frac{1}{(\mu+d+\delta+\frac{a}{b})}$ represents the average life expectancy of infectious individuals, $\frac{A}{\mu}$ represents the number of susceptible individuals at the beginning of the infectious process and $\frac{\beta A}{\mu}$ represents the value when all individuals are susceptible. Hence, our basic reproduction number R_0 is biologically well defined.

A. Analysis for $R_0 \neq 1$

Clearly, Eq. (10) always has one negative root $\lambda_1 = -\mu$ and one other root is determined by the solution of equation

$$\lambda + \mu + d + \delta + \frac{a}{b} - \frac{\beta A}{\mu}e^{-\lambda\tau} = 0.$$

Let

$$f(\lambda) = \lambda + \mu + d + \delta + \frac{a}{b} - \frac{\beta A}{\mu}e^{-\lambda\tau}.$$

If $R_0 > 1$, it is readily seen that for λ real,

$$f(0) = \mu + d + \delta + \frac{a}{b} - \frac{\beta A}{\mu} = \left(\mu + d + \delta + \frac{a}{b}\right)(1 - R_0) < 0, \lim_{\lambda\to+\infty} f(\lambda) = +\infty.$$

Hence, $f(\lambda) = 0$ has a positive real root if $R_0 > 1$.

If $R_0 < 1$, we assume that $Re\lambda \geq 0$.

We notice that

$\mathrm{Re}\lambda = \frac{\beta A}{\mu} e^{-\mathrm{Re}\lambda\tau} \cos \mathrm{Im}\lambda\tau - \left(\mu + d + \delta + \frac{a}{b}\right) < \frac{\beta A}{\mu} - \left(\mu + d + \delta + \frac{a}{b}\right) = \left(\mu + d + \delta + \frac{a}{b}\right)(R_0 - 1) < 0$. a contradiction to our assumption. Hence, if $R_0 < 1$ then λ is a root of Eq. (10) with negative real part.

Hence, we state the following theorem [3]:

Theorem 2 *The disease-free equilibrium* $Q = Q(\frac{A}{\mu}, 0)$ *is locally asymptotically stable if* $R_0 < 1$ *and unstable if* $R_0 > 1$.

B. Analysis at $\mathbf{R_0 = 1}$

We notice that the system (5)–(6) evaluated at $R_0 = 1$ and $\beta = \beta^* = \frac{\mu(\mu+d+\delta+\frac{a}{b})}{A}$ has a zero eigenvalue and another eigenvalue is negative. The stability behavior of equilibrium point at $R_0 = 1$ cannot be determined using linearization so we use Center manifold theory [16]. In order to apply center manifold theorem to model, we redefined $S = x_1$ and $I = x_2$ then the system (5)–(6) can be rewritten as

$$\frac{dx_1}{dt} = A - \mu\, x_1 - \frac{\beta\, x_1\,(t-\tau)\, x_2\,(t-\tau)}{1 + \alpha\, x_2\,(t-\tau)} \equiv f_1, \tag{11}$$

$$\frac{dx_2}{dt} = \frac{\beta\, x_1(t-\tau) x_2(t-\tau)}{1 + \alpha\, x_2(t-\tau)} - (\mu + d + \delta)\, x_2 - \frac{a x_2}{b + x_2^2} \equiv f_2. \tag{12}$$

Let J^* be the Jacobian matrix at $R_0 = 1$ and bifurcation parameter $\beta = \beta^*$. Then

$$J^* = \begin{bmatrix} -\mu & -\frac{\beta^* A}{\mu} \\ 0 & 0 \end{bmatrix}.$$

Let $u = [u_1, u_2]$ and $w = [w_1, w_2]^T$ be the left eigenvector and right eigenvector of J^* corresponding to the zero eigenvalue. Then we have

$$u_1 = 0, u_2 = 1 \text{and } w_1 = -\frac{\beta^* A}{\mu^2}, w_2 = 1.$$

The non-zero partial derivatives associated with the functions of the system (11)–(12) Evaluated at $R_0 = 1$ and $\beta = \beta^*$ are

$$\left(\frac{\partial^2 f_2}{\partial x_1 \partial x_2}\right)_Q = \beta^*, \left(\frac{\partial^2 f_2}{\partial x_2 \partial x_1}\right)_Q = \beta^*, \left(\frac{\partial^2 f_2}{\partial x_2^2}\right)_Q = \frac{-2\alpha\beta^* A}{\mu} \text{and } \left(\frac{\partial^2 f_2}{\partial x_2 \partial \beta^*}\right)_Q = \frac{A}{\mu}.$$

Then from Theorem 4.1 of Casillo–Chavez and Song [6], the bifurcation constants a_1 and b_1 are

$$\begin{aligned} a_1 &= \sum_{k,i,j=1}^{2} u_k w_i w_j \left(\frac{\partial^2 f_k}{\partial x_i \partial x_j}\right)_Q, \\ &= u_2 \left(2 w_1 w_2 \beta^* + w_2^2 \left(\frac{-2\alpha\beta^* A}{\mu}\right)\right), \\ &= -\left(\frac{2\beta^* A(\beta^* + \mu\alpha)}{\mu^2}\right) < 0, \end{aligned}$$

and

$$b_1 = \sum_{k,i=1}^{2} u_k w_i \left(\frac{\partial^2 f_k}{\partial x_i \partial \beta^*}\right)_Q,$$

$$= u_2\left(w_2\frac{A}{\mu}\right),$$
$$= \frac{A}{\mu} > 0.$$

Thus from Theorem 4.1(iv) of Casillo–Chavez and Song [6], we state the following theorem:

Theorem 3 *The disease-free equilibrium* $Q = Q\left(\frac{A}{\mu}, 0\right)$ *changes its stability from stable to unstable at* $R_0 = 1$ *and there exists a positive equilibrium as* R_0 *crosses one. Hence, the system* (11)–(12) *undergoes a transcritical bifurcation at* $R_0 = 1$.

Endemic Equilibrium

To find the conditions for the existence of an equilibrium $Q^* = (S^*, I^*)$ for which the disease is endemic in the population, the system (5)–(6) is rearranged to get S^*, and I^* which gives

$$\left(S^* = \frac{\left(1+\alpha I^*\right)\left((\mu+d+\delta)\left(b+I^{*2}\right)+a\right)}{\beta\left(b+I^{*2}\right)}, I^*\right)$$

Here I^* is given by the following equation

$$C_1 I^{*3} + C_2 I^{*2} + C_3 I^* + C_4 = 0 \tag{13}$$

where

$$C_1 = (\mu+d+\delta)(\mu\alpha+\beta),\ C_2 = \mu(\mu+d+\delta) - A\beta,\ C_3 = (b(\mu+d+\delta)+a)(\mu\alpha+\beta),$$
$$C_4 = -\beta bA + \mu a + \mu^2 b + \delta b + db.$$

Now using Descartes's rule of signs [19], the cubic equation has a unique positive real root I^* if the following condition hold true:

$$C_1 > 0, C_2 > 0, C_3 > 0 and C_4 < 0.$$

After finding the value of I^* we can find the value of S^*. This implies that there exists an endemic equilibrium $Q^*(S^*, I^*)$ if above condition hold true.

The local stability of Q^* is explored as follows: the characteristic equation of the system (5)–(6) evaluated at endemic equilibrium point Q^* is a second degree transcendental polynomial which is

$$\lambda^2 + p_0\lambda + q_0 + \left(p_1\lambda + q_1\right)e^{-\lambda\tau} = 0,$$

where

$$p_0 = \left((2\mu+d+\delta) + \frac{a\left(b-I^{*2}\right)}{\left(b+I^{*2}\right)^2}\right),\ q_0 = \mu\left((\mu+d+\delta) + \frac{a\left(b-I^{*2}\right)}{\left(b+I^{*2}\right)^2}\right),$$
$$p_1 = \left(\frac{\beta I^*}{(1+\alpha I^*)} - \frac{\beta S^*}{(1+\alpha I^*)^2}\right),\ q_1 = \left(\beta\left((\mu+d+\delta) + \frac{a\left(b-I^{*2}\right)}{\left(b+I^{*2}\right)^2}\right)\frac{I^*}{(1+\alpha I^*)} - \frac{\mu\beta S^*}{(1+\alpha I^*)^2}\right)$$

Theorem 4 *For* $\tau = 0$, Q^* *is locally asymptotically stable if both conditions* $\frac{S^*}{(1+\alpha I^*)I^*} \leq 1$, and $\frac{\mu S^*}{(1+\alpha I^*)} \leq \left((\mu+d+\delta) + \frac{a\left(b-I^{*2}\right)}{\left(b+I^{*2}\right)^2}\right) I^*$ *are satisfied simultaneously.*

Proof At Q^*, the characteristic equation at $\tau = 0$ is given by

$$\lambda^2 + p_0\lambda + q_0 + (p_1\lambda + q_1) = 0. \tag{14}$$

It is easy to show that if both conditions $\frac{S^*}{(1+\alpha I^*)I^*} \leq 1$, and $\frac{\mu S^*}{(1+\alpha I^*)} \leq \left((\mu + d + \delta) + \frac{a(b-I^{*2})}{(b+I^{*2})^2}\right) I^*$ are satisfied simultaneously.

$$\begin{aligned}
p_0 + p_1 &= \left((2\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \left(\frac{\beta I^*}{(1 + \alpha I^*)} - \frac{\beta S^*}{(1 + \alpha I^*)^2}\right) \\
&= \left(\mu + (\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \frac{\beta}{(1 + \alpha I^*)}\left(I^* - \frac{S^*}{(1 + \alpha I^*)}\right) > 0, \\
q_0 + q_1 &= \mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \left(\beta\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)\frac{I^*}{(1 + \alpha I^*)} - \frac{\mu\beta S^*}{(1 + \alpha I^*)^2}\right) \\
&= \mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \frac{\beta}{(1 + \alpha I^*)}\left(\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) I^* - \frac{\mu S^*}{(1 + \alpha I^*)}\right) > 0.
\end{aligned}$$

Hence, using Routh–Hurwitz criterion, the endemic equilibrium Q^* of the model is locally asymptotically stable at $\tau = 0$.

Theorem 5 *For $\tau > 0$, Q^* is locally asymptotically stable if all three conditions $\frac{\mu S^*}{(1+\alpha I^*)} \leq \left((\mu + d + \delta) + \frac{a(b-I^{*2})}{(b+I^{*2})^2}\right) I^*$, $\frac{\beta}{\mu} \leq \frac{(1+\alpha I^*)}{I^*}$ and $M_2 < M_1$ are satisfied simultaneously,where*

$$\begin{aligned}
M_1 &= \mu^2 + \left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)^2, \\
M_2 &= \left(\frac{\beta I^*}{(1 + \alpha I^*)} - \frac{\beta S^*}{(1 + \alpha I^*)^2}\right)^2.
\end{aligned}$$

Proof At Q^* the characteristic equation at $\tau > 0$ is given by

$$\lambda^2 + p_0\lambda + q_0 + (p_1\lambda + q_1)\, e^{-\lambda\tau} = 0. \tag{15}$$

For $\tau > 0$, by Corollary 2.4 in Ruan and Wei [15], it follows that if instability occurs for a particular value of the delay τ, a characteristic root of (15) must intersect the imaginary axis. Therefore, we take $\lambda = i\omega$, $\omega > 0$ is the roots of the characteristic Eq. (15). Substituting $\lambda = i\omega$ in the Eq. (15), we get

$$-\omega^2 + q_0 + p_1\omega \mathrm{Sin}\omega\tau + q_1 \mathrm{Cos}\omega\tau + i\left(p_1\omega \mathrm{Cos}\omega\tau - q_1 \mathrm{Sin}\omega\tau + p_0\omega\right) = 0. \tag{16}$$

On separating real and imaginary part of (10)

$$p_1\omega \mathrm{Sin}\omega\tau + q_1 \mathrm{Cos}\omega\tau = \omega^2 - q_0, \tag{17}$$

$$p_1\omega \mathrm{Cos}\omega\tau - q_1 \mathrm{Sin}\omega\tau = -p_0\omega. \tag{18}$$

On squaring and adding the both side of Eqs. (17) and (18) yield

$$\omega^4 + \left(p_0^2 - 2q_0 - p_1^2\right)\omega^2 + \left(q_0^2 - q_1^2\right) = 0. \tag{19}$$

Letting $\omega^2 = z_1$, Eq. (19) becomes

$$z_1^2 + Pz_1 + T = 0, \tag{20}$$

where $P = \left(p_0^2 - 2q_0 - p_1^2\right)$ and $T = \left(q_0^2 - q_1^2\right)$

It is easy to show that if all three conditions $\frac{\mu\, S^*}{(1+\alpha\, I^*)} \leq \left((\mu + d + \delta) + \frac{a(b-I^{*2})}{(b+I^{*2})^2}\right) I^*$, $\frac{\beta}{\mu} \leq \frac{(1+\alpha\, I^*)}{I^*}$ and $M_2 < M_1$ are satisfied simultaneously,

$$
\begin{aligned}
P = \left(p_0^2 - 2q_0 - p_1^2\right) &= \left((2\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)^2 \\
&\quad - 2\mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) - \left(\frac{\beta\, I^*}{(1 + \alpha\, I^*)} - \frac{\beta\, S^*}{(1 + \alpha\, I^*)^2}\right)^2 \\
&= \mu^2 + \left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)^2 - \left(\frac{\beta\, I^*}{(1 + \alpha\, I^*)} - \frac{\beta\, S^*}{(1 + \alpha\, I^*)^2}\right)^2 \\
&= M_1 - M_2 > 0,
\end{aligned}
$$

$$
\begin{aligned}
T = \left(q_0^2 - q_1^2\right) &= \left(q_0 - q_1\right)\left(q_0 + q_1\right), \\
&= \left(\mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) - \left(\beta\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)\frac{I^*}{(1 + \alpha\, I^*)} - \frac{\mu\beta S^*}{(1 + \alpha\, I^*)^2}\right)\right) \\
&\quad \times \left(\mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \left(\beta\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)\frac{I^*}{(1 + \alpha\, I^*)} - \frac{\mu\beta S^*}{(1 + \alpha\, I^*)^2}\right)\right) \\
&= \left(\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)\left(\mu - \frac{\beta\, I^*}{(1 + \alpha\, I^*)}\right) + \frac{\mu\beta S^*}{(1 + \alpha\, I^*)^2}\right) \\
&\quad \times \left(\mu\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right) + \left(\beta\left((\mu + d + \delta) + \frac{a\left(b - I^{*2}\right)}{\left(b + I^{*2}\right)^2}\right)\frac{I^*}{(1 + \alpha\, I^*)} - \frac{\mu\beta S^*}{(1 + \alpha\, I^*)^2}\right)\right) > 0.
\end{aligned}
$$

By Routh–Hurwitz Criterion, this contradicts to our assumption for instability i.e. $\lambda = i\omega$. Hence, it is proved that the endemic equilibrium Q^* of the model is locally asymptotically stable for $\tau > 0$.

Numerical Simulation and Discussion

In this section, we simulate the system (5)–(6). All computational work has been carried out by MATLAB 2012b. We choose the following set of tested values of the parameters from the literature [2, 4, 11] (Table 2).

Table 2 List of parameters

Parameter	Value	Unit
A	5	person $(\text{day})^{-1}$
α	0.002	$(\text{person})^{-1}\,(\text{day})^{-1}$
β	0.004	$(\text{person})^{-1}\,(\text{day})^{-1}$
μ	0.05	$(\text{day})^{-1}$
d	0.001	$(\text{day})^{-1}$
δ	0.002	$(\text{day})^{-1}$
a	0.2	$(\text{person})^{-1}\,(\text{day})^{-1}$
b	0.004	$(\text{person})^{-1}\,(\text{day})^{-1}$

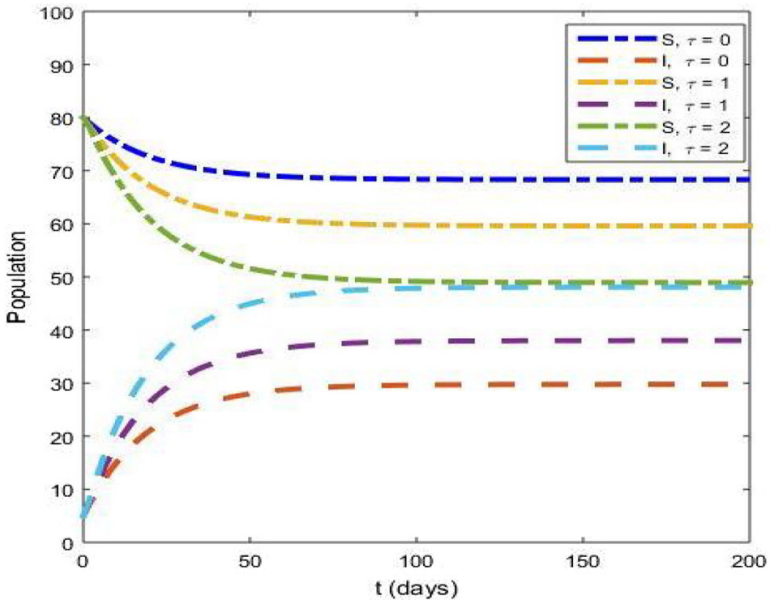


Fig. 2 Susceptible and Infected populations at various values of time lag τ

Initial values are taken as

$$S(0) = 80, I(0) = 5.$$

Figure 2 shows the combined population of susceptible and infected at time lag $\tau = 0$, 1, and 2days. Clearly, as the time delay is increasing the number of susceptible is decreasing and the number of infected is increasing. Furthermore, both populations are decreasing and increasing respectively to attain a steady state. Hence, it can be concluded that the time delay plays an important role to understand the infection progression in human population.

Figure 3 shows the infected population at various values of transmission rate β. Clearly, it can be seen that as the values of β increases the numbers of infected individuals also increase. Furthermore, the infected population attains its steady state because of treatment.

Figure 4 shows the infected population at various values of inhibition rate taken by infected α. Clearly, infected population is decreasing with the increase in the value of α.

Figure 5 shows the infected population at various values of treatment rate a. Clearly, as the value of treatment rate is increasing the infected population is decreasing.

Figure 6 shows the infected population with and without Monod–Haldane type treatment rate. Clearly, the infected population is less with Monod–Haldane type treatment rate in the comparison to the infected population without Monod–Haldane type treatment rate. Hence, Monod–Haldane function type treatment rate play a crucial role in controlling the infection in the population.

Conclusion

In the present article, we proposed a delayed susceptible–infected–recovered (S–I–R) model along with an incidence rate as Holling type II and treatment rate as Monod–Haldane type. We showed that the model has two equilibrium points: disease-free and endemic equilibrium points. We showed the stability of the model for disease-free equilibrium with the help of basic reproduction number R_0. We proved that the disease-free equilibrium of the model is locally asymptotically stable when $R_0 < 1$ and unstable when $R_0 > 1$ for time lag $\tau \geq 0$.

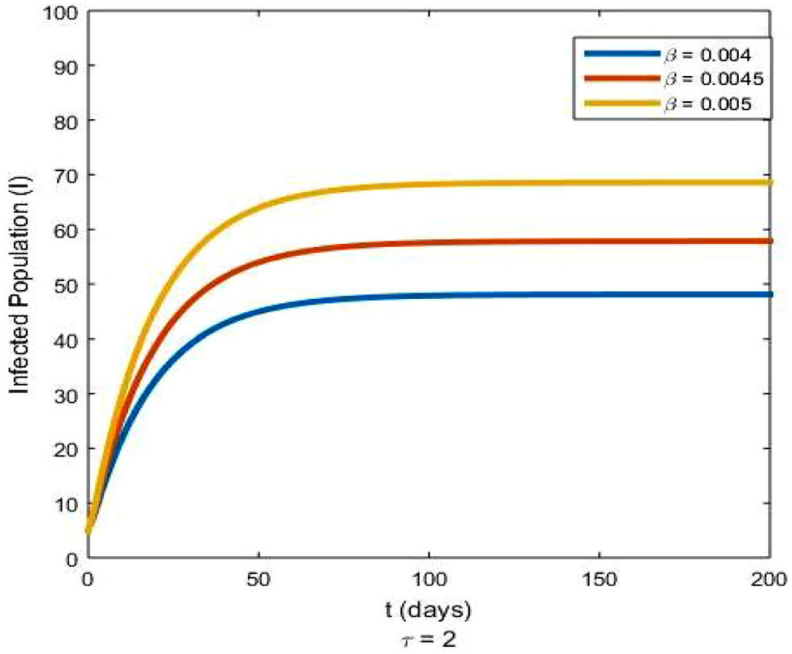


Fig. 3 Infected population at various values of transmission rate **β**

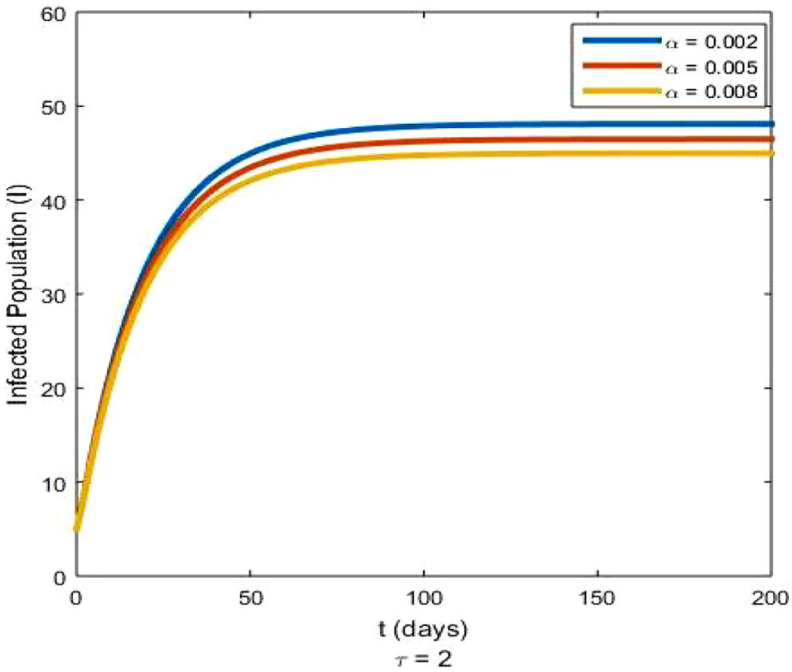


Fig. 4 Infected population at various values of measure of inhibition **α**

Clearly, it indicates that the infection persists along with the low availability of treatment when the basic reproduction number is greater than one. The local stability of the model for disease-free equilibrium at $R_0 = 1$ has been discussed by center manifold theory and it is investigated that the model change its stability from stable to unstable at $R_0 = 1$ for disease-free equilibrium and there also exists a positive equilibrium as R_0 crosses one. Hence the model undergoes a transcritical bifurcation at $R_0 = 1$ for time lag $\tau \geq 0$.The stability of the endemic equilibrium of the model has been investigated by Routh–Hurwitz criterion. We investigated that endemic equilibrium point of the model is locally asymptotically stable when the Theorems 4 and 5 hold true for time lag $\tau = 0$ and $\tau > 0$ respectively. Simulation

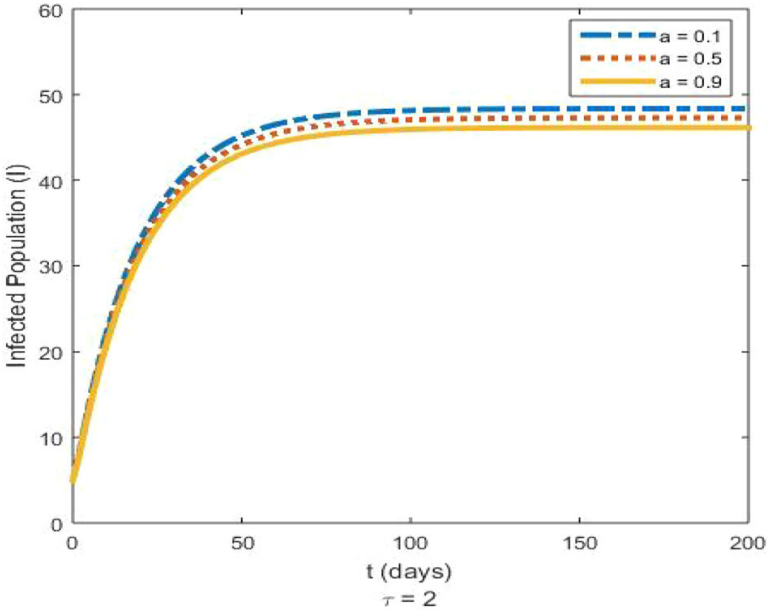


Fig. 5 Infected population at various values of treatment rate $'a'$

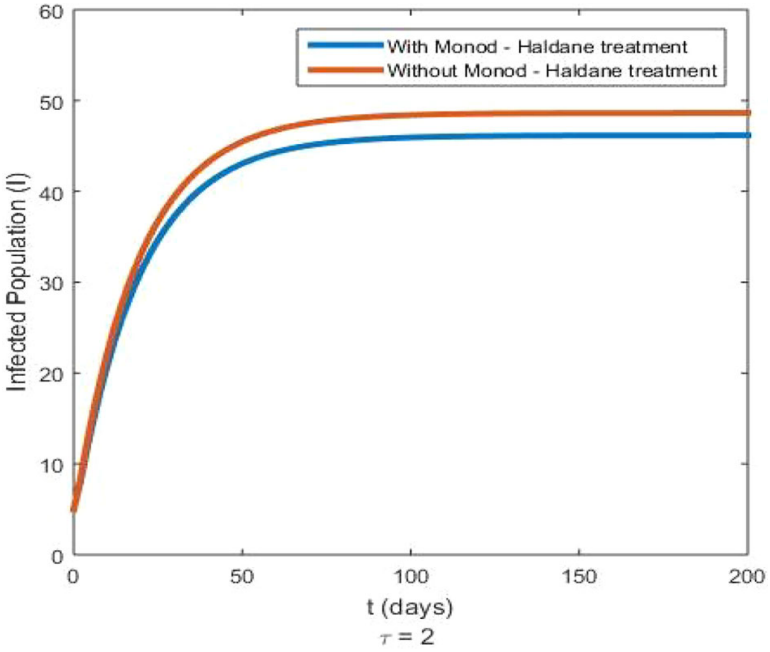


Fig. 6 Infected population with and without Monod–Haldane type treatment rate

has been carried out with the help of MATLAB 2012b to observe the effects of time delay and Monod–Haldane type treatment rate. The numerical simulation indicates that the infection will increase with the increased in transmission rate, but settles down at a steady state due to the availability of treatment facilities. Further, infection will decrease with the increase in measure of inhibition taken by infected. With the help of numerical simulations, we also found that the eradication of the infection is possible only when the treatment given to the population is properly managed according to the availability of the resources.

Acknowledgements The authors would like to gratefully acknowledge Delhi Technological University, Delhi, India for providing financial support to carry out this research work. Authors are also grateful to the anonymous referees for their valuable reviews and suggestions which improved the quality of the paper.

References

1. Gumel, A.B., Connell Mccluskey, C., Watmough, J.: An SVEIR model for assessing the potential impact of an imperfect anti-SARS vaccine. Math. Biosci. Eng. **3**, 485–494 (2006)
2. Kumar, A., Nilam, : Stability of a time delayed SIR epidemic model along with nonlinear incidence rate and Holling type-II treatment rate. Int. J. Comput. Methods **15**(1), 1850055 (2018)
3. Dubey, B., Patara, A., Srivastava, P.K., Dubey, U.S.: Modelling and analysis of a SEIR model with different types of nonlinear treatment rates. J. Biol. Syst. **21**(3), 1350023 (2013)
4. Dubey, B., Dubey, P., Dubey, Uma S.: Dynamics of a SIR model with nonlinear incidence rate and treatment rate. Appl. Appl. Maths. **2**(2), 718–737 (2016)
5. Capasso, V., Serio, G.: A generalization of the Kermack–Mckendrick deterministic epidemic model. Math. Biosci. **42**, 41–61 (1978)
6. Castillo-Chavez, C., Song, B.: Dynamical models of tuberculosis and their applications. Math. Biosci. Eng. **1**, 361–404 (2004)
7. Baek, H., Do, Y., Saito, Y.: Analysis of an impulsive predator–prey system with Monod-Haldane functional response and seasonal effects. Math. Prob. Eng. **29**, Article Id 543187 (2009)
8. Andrews, J.F.: A mathematical model for the continuous culture of microorganisms untilizing inhibitory substrates. Biotechnol. Bioeng. **10**, 707–723 (1968)
9. Hattaf, K., Yousfi, N.: Mathematical model of influenza A (H1N1) infection. Adv. Std. Biol. **1**, 383–390 (2009)
10. Hattaf, K., Lashari, A.A., Louartassi, Y., Yousfi, N.: A delayed SIR epidemic model with general incidence rate. Electron. J. Qual. Theory Differ. Equ. **3**, 1–9 (2013)
11. Dubey, P., Dubey, B., Dubey, U.S.: An SIR model with nonlinear incidence rate and Holling type III treatment rate. Appl. Anal. Biol. Phys. Sci. **186**, 63–81 (2015)
12. Driessche, P.V.D., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartment models of disease transmission, math. Bioscience **180**, 29–48 (2002)
13. Xu, R., Ma, Z.: Stability of a delayed SIRS epidemic model with a nonlinear incidence rate. Chaos Solut. Fract. **41**, 2319–2325 (2009)
14. Sahani, S.K., Yashi: Analysis of a delayed HIV infection model, International workshop on a computational intelligence (IWCI) Dhaka, Bangladesh, pp. 246–251. https://doi.org/10.1109/iwci.2016.7860374 (2016)
15. Ruan, S., Wei, J.: On the zeros of transcendental functions with applications to stability of delay differential equations with two. Dyn. Contin. Discret. Impuls. Syst. Ser. A Math. Anal. **10**, 863–874 (2003)
16. Sastry, S.: Analysis, Stability and Control. Springer, New York (1999)
17. Tipsri, S., Chinviriyasit, W.: Stability analysis of SEIR model with saturated incidence and time delay. Int. J. Appl. Phys. Math. **4**(1), 42–45 (2014)
18. Sokol, W., Howell, J.A.: Kineties of phenol oxidation by ashed cell. Biotechnol. Bioeng. **23**, 2039–2049 (1980)
19. Wang, X.: A simple proof of Descartes's rule of signs. Am. Math. Mon. (2004). https://doi.org/10.2307/4145072

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/324941227

Effect of Non-linearity on Wheel/rail Interaction Dynamics using Bond Graph

Conference Paper · January 2018

CITATIONS
0

READS
11

3 authors:



Vivek Kumar
Sant Longowal Institute of Engineering and Technology
13 PUBLICATIONS **13** CITATIONS
SEE PROFILE



Rastogi Vikas
Delhi Technological University
66 PUBLICATIONS **174** CITATIONS
SEE PROFILE



Pushparaj Mani Pathak
Indian Institute of Technology Roorkee
121 PUBLICATIONS **482** CITATIONS
SEE PROFILE

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

Project Modelling and control of mobile cooperating bionic arms View project

Project Modeling and Control of Quadruped Robot with Flexible Legs View project

All content following this page was uploaded by Vivek Kumar on 04 May 2018.

The user has requested enhancement of the downloaded file.

1st International Conference on New Frontiers in Engineering, Science & Technology,
New Delhi, India, January 8-12, 2018

NFEST/2018/R–180

Effect of Non-linearity on Wheel/rail Interaction Dynamics using Bond Graph

Vivek Kumar[1], Vikas Rastogi[2], P.M. Pathak[3]

[1]*Department of Mechanical Engineering, Sant Longowal Institute of Engineering and Technology, Sangrur, India*
[2]*Department of Mechanical, Production, Industrial and Automobile Engineering, Delhi Technological University, Delhi, India*
[3]*Department of Mechanical and Industrial Engineering, India Institute of Technology, Roorkee, India*
[1]*vivku@rediffmail.com*

Abstract: *The way railway vehicles interact with the track has been the subject of interest for railway industries from several decades. With the increasing demands of speed and load carrying capacity, there is a growing concern about the track system, its dynamics along with specific attention towards wheel/track interaction dynamics in order to improve vehicle/track design to reduce track and vehicle maintenance and assure safety. Thus, a train must be designed to guarantee safe operation by identifying and eliminating the causes of the derailment. Today'strains are complex mechanical systems that can be better analyzed and designed using modern modeling and simulation techniques. Development of a detailed model of railway vehicles system is preliminary requirement to study vehicle performance, to improve existing designs, to develop new ones, and to develop safety procedures for different working and loading conditions. Computer simulations play an important role in the evaluation of railway vehicle performance [1].*

A large variety of models has been proposed and is used for many different circumstances. The accurate of prediction of interaction force between the rail and wheel is of great importance to the rail industry for detecting the potential causes responsible for vehicle and track damage. In present work, a combined vehicle-track bond graph model is developed to study the wheel/ track impact. The rail flexible is considered modeling rail as a flexible Euler Bernoulli beam with discrete support. Non-linearity in pad and ballast behavior and wheel-rail contact is also taken into considerations while developing the combined model. It is noticed from *simulation results that; a linear track model is found to be inappropriate for wheel/track interaction problems due to the extremely large value of impact load. The obtained results will further lead to provide a good all-inclusive view of the contact dynamics of the wheel/track interface.*

Keywords: *Wheel-track interaction; Bond graph model; Non-linear Track; Contact force; Simulation; wheel flat.*

1. INTRODUCTION

The energy exchange between a train and track plays an important role in the estimation of wheel/rail interaction force due to wheel flat and other wheel/rail irregularities. Wheel flats were developed due locking of thewheel while breaking. Large wheel flats caused safety hazards whereas smaller contributes to track/vehicle damage and increased maintenance cost. The impact forces generated due to rail/wheel defect can be extremely large and may cause serious failures in the vehicle or track components.

This situation can cause the wheel-sets to impact on to the track, resulting in high impact force and a forced vibration with high frequencies. These high impact forces may lead to damage to track components. The impact force further increases with running speed of the vehicle. Therefore, the accurate prediction of wheel-rail impact forces from theview of vehicle/track design and maintenance is of paramount concern and is critical also for thesafe and comfortable running of the trains. These days railway companies around the world are under constant pressure to compete with air and road transport system. Worldwide attention, therefore, has been directed towards improving the efficiency of railway vehicle in terms of speed, safety and comfort [2].

The vertical dynamics of railway vehicle/track system at high speed is mainly dominated by wheel/rail interaction and dynamics of wheel and track. The most important feature in the modeling of track structure is the realistic representation of its characteristics, irregularities and defects present on track. On the other hand, vehicle model can be simplified without losing accuracy, as long as wheel/ rail contact model is accurate.

Thus, no. Of experimental and analytical studies have been conducted to study the train/track interaction in the presence of wheel/rail defects. The literature related to vehicle/track interaction dynamics has been reviewed by many researchers and few good surveys have been published. Kerr [3] and Fryba [4] have given historical

surveys and solution method review on the steady-state response of infinite beams subjected to moving load. Knothe and Grassie [5] have presented a detailed review on track and vehicle/track interaction. Newton and Clark [6] have presented the experimental and analytical study to measure the impact forces during the passage of train from the track with wheel flat. Fermer and Nielsen [7] have conducted a full-scale experiment using a wagon loaded with instrumented wheelsets. Johansson and Nielsen [8] have studied the effect of out-of-round wheels on the wheel-rail interaction force through experimentation and simulation. Sun and Dhaneskar [9] have given a vertical dynamic model for analysis of the track and wagon system. Uzzal et. Al. [10] have used a 5 DOF pitch plane vehicle model to study the influence of wheel flats on the rail/ wheel interaction force. A multi-point rail-wheel contact model is proposed by Dong et al [10] to study the wheel/rail interaction analysis. Zhai and Cai [11] have reported a 10 dof vehicle model and rail as beam supported discretely at rail/sleeper junction to study the vertical dynamic interaction.

Dahlberg [12] has given a non-linear model of the track for the analysis of the train track interaction. Wu and Thompson [13] have presented wheel/track interaction analysis using a nonlinear track model. Though the elastic behavior of these components is non-linear, then also, they are modelled in most of studies discussed above through linear relation. Nevertheless, this may lead to improper results in certain circumstance due to variation in the pad and ballast stiffness with loads.

In this work wheel/rail interaction behavior is investigated, considering a nonlinear track model. The nonlinear elastic behaviour of pad and ballast components is modelled through a nonlinear relation with corresponding displacement component. In the present work bond graph formulism is used to study the wheel/rail interaction dynamics. In bond graph modeling a topological structure is used to represent the energy characteristics of engineering systems. Bond graphs are a concise pictorial representation of different interacting energy domains, several authors use this methodology as a modeling tool to study different aspects of dynamics of rail, and road vehicle [14-21].

2. RAIL VEHICLE /TRACK SYSTEM MODEL

The analysis of wheel-rail impact force or acceleration responses in the presence of wheel/rail defects requires development of a comprehensive vehicle and track system model, wheel-rail contact model along with wheel/rail irregularity model. A coupled wheel/track interaction model is used and is shown in Figure 1. A single degree of vehicle model coupled with a track model is used to study the effects of wheel flat on the magnitudes of the impact forces. The schematic of the coupled vehicle/track model is shown in Figure 1. The vehicle is modelled as wheel that represents the unsprung mass acted by static load P0. The track system consists of rail, pads, sleepers, ballasts and subgrade. The rail is considered as Euler-Bernoulli's beam discretely supported on the sleepers through the rail pads and fasteners. The rail pads and fasteners are assumed as massless spring-damper element inserted between the rail and sleepers. It is assumed that sleepers are equivalent to lumped masses, which are upheld on the ballast mass. The ballasts are assumed as massless spring-damper combination inserted between each sleeper and subgrade.

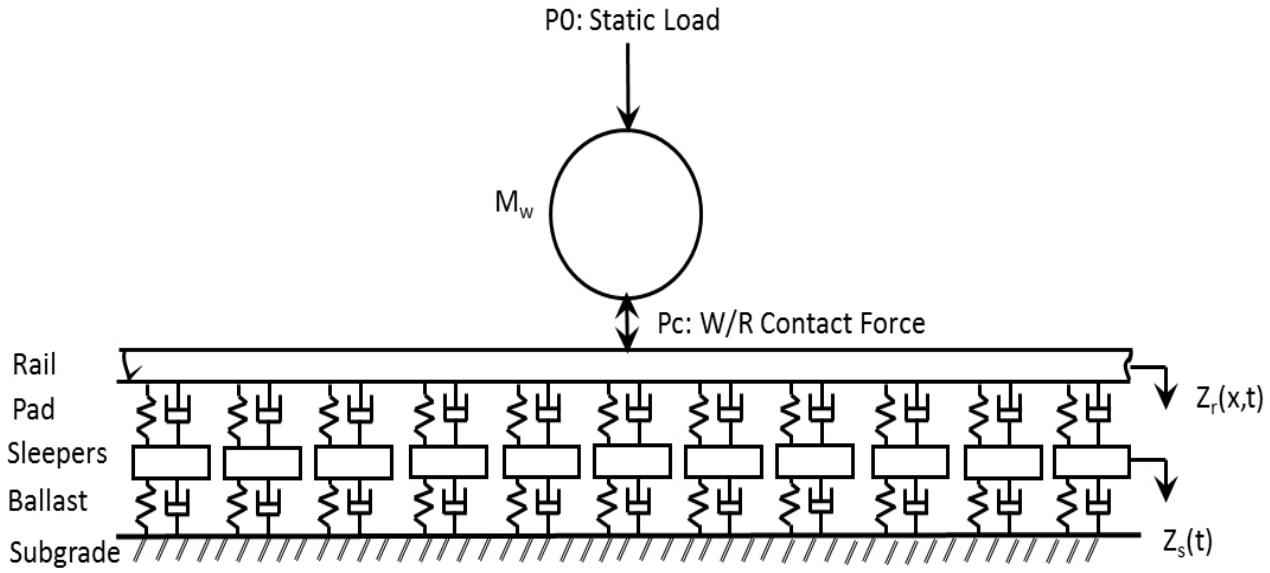


Fig. 1. Coupled rail vehicle/track model

A building block approach is adopted in which each component is modeled separately and then all components are combined to construct an integrated model of the vehicle/track system. Figure 2 shows the bond graph model of railway vehicle and is termed as RV model.

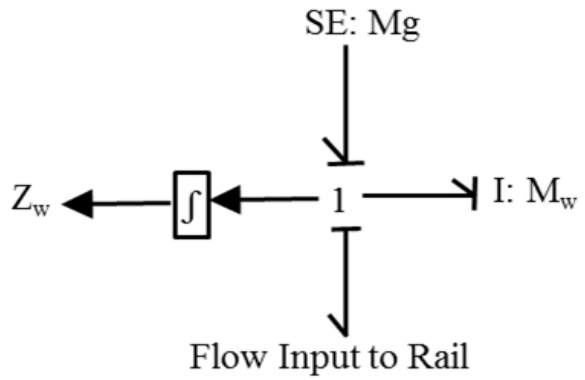


Fig. 2. Bond graph of vehicle sub model

The flexibility of rail is considered by assuming it as eulerbernoulli's beam with flexural rigidity EI.The transverse deflection of the rail being expressed as $z_r(x, t)$. The partial differential equation of the beam can be derived and given by

$$EI\frac{\partial^4 z_r(x,t)}{\partial x^4} + \rho A\frac{\partial^2 z_r(x,t)}{\partial x^2} = P_c(t)\delta(x - x_w) - \sum_{j=1}^{NS} F_{rsj}\delta(x - x_{sj}) \quad (1)$$

Where, ρ is density, A is area of cross section of rail, $P_c(t)$ is the force at the wheel/ rail contact, F_{rsj} is the force developed at the j^{th} sleeper/rail interface, x_w Is the position of the wheel at any time t, and x_{sj} Is the position coordinates of j^{th}sleeper, NS is no. Of sleepers and $\delta(.)$ Is dirac's delta function.

The method of separation is used to convert partial differential equation into ordinary derivatives. This method providesolution as an infinite sum of product of shape function $Y_i(x)$ and modal amplitude $q_i(t)$. Assuming the ends of rail to be hinged shape function is given by Eq. 2. In the present analysis total, 70 modes have been considered.

$$Y_i(x) = \mathrm{Sin}\left(\frac{i\pi x}{L}\right) \quad (2)$$

The deflection of rail in the vicinity of the contact points is given as

$$z_r(\mathrm{X},\mathrm{t}) = \sum_{\mathrm{I}=1}^{\mathrm{NM}} \mathrm{Y_I(X)Q_I(t)} \quad (3)$$

Integrating Eq. (1) with respect to x after substituting thevalue of mode shape Y_i and rail deflection from Eqs. (2) -(3) and considering the orthogonality property of shape functions, one gets Eq. (4)

$$m_i\ddot{q}_i + k_i q_i = P_c(t)Y_i(x_w) - \textstyle\sum_{i=1}^{NS} F_{rsj}Y_i(x_{sj}) \quad (4)$$

Where, m_i is the modal mass and k_i Is the i^{th} mode stiffness and are given by Eq. (5)

$$m_i = \frac{\rho AL}{2}, \quad k_i = m_i\omega_i^2, \qquad i = 1, 2 \ldots . NM \quad (5)$$

3. NON-LINEAR WHEEL/RAIL INTERACTION

The interaction of the wheel with the rail differs it with other means of land transport. The interaction force between wheel and rail is considered in the model through Hertzian nonlinear theory. According to Hertzian theory,

$$P_c(t) = K_H(\Delta z)^{\frac{3}{2}} \quad (6)$$

This force becomes zero if the wheel loses contact with the rail i.e. When $\Delta z \leq 0$

Where, Δz is the wheel-rail overlap and is given by

$$\Delta z(t) = z_w - z_r - z_{irr} \quad (7)$$

Where, z_w, z_r Are wheel and rail displacement respectively and z_{irr}Is the wheel flat function.

The non-linear elastic behaviour of pad and ballast can be expressed by Eqs. (8) -(11) [15].

$$K_{pad} = 52 + 6.24 \times 10^8 x_p^2 \; MN/m \quad (8)$$

$$K_{ballast} = 22.75 + 2.6 \times 10^8 x_b^2 \; MN/m \quad (9)$$

$$R_{pad} = \eta_1\sqrt{m_1 K_{pad}} \quad (10)$$

$$R_{ballast} = \eta_2\sqrt{m_2 K_{ballst}} \quad (11)$$

Where, x_p And x_b Are pad and ballast compression respectively in meters, $\eta_1 = 0.25$, and $\eta_2 = 1$ are the loss factor, $m_1$and $m_2$are equivalent mass and are given by Eq. (12).

$$m_1 = \frac{m_s m_r}{m_s + m_r}, \qquad m_2 = m_s + m_r \quad (12)$$

Here, m_s And m_r Are the sleeper mass and rail mass in sleeper bay respectively in kg.

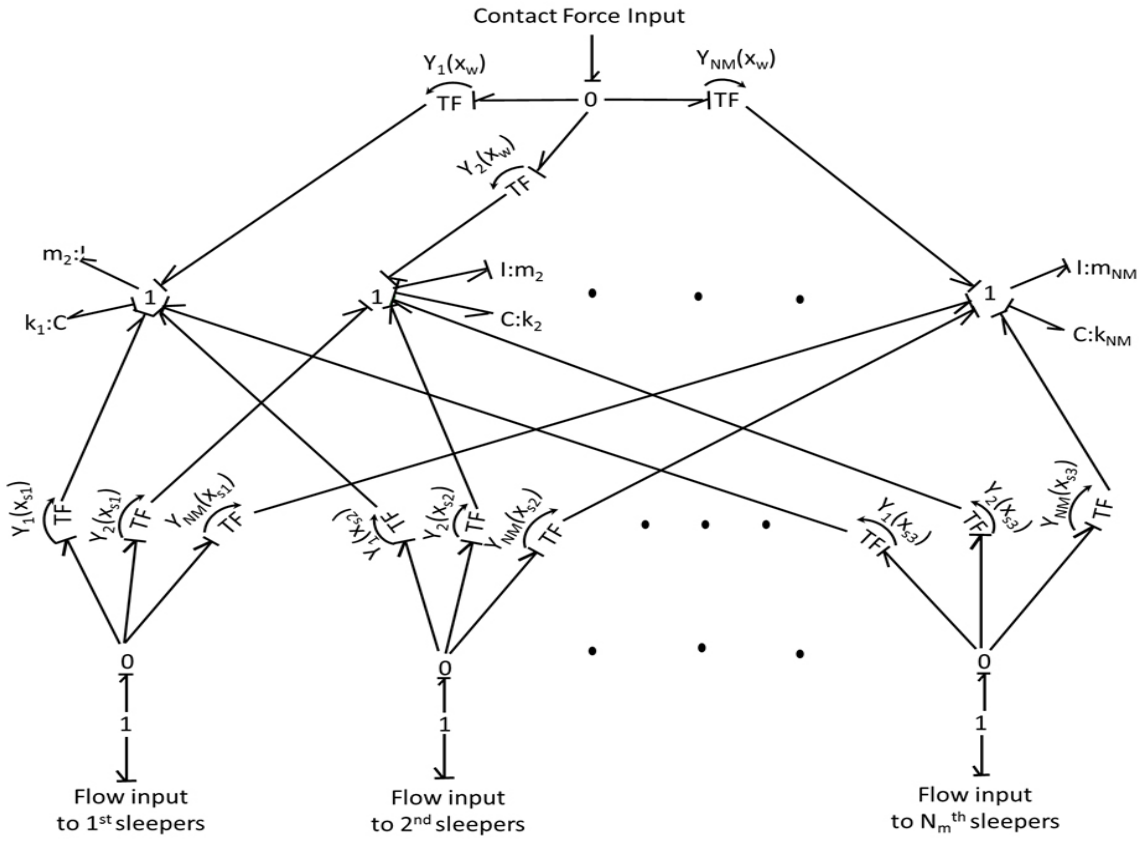


Fig. 3. Bond graph representation of rail beam (RB) sub model

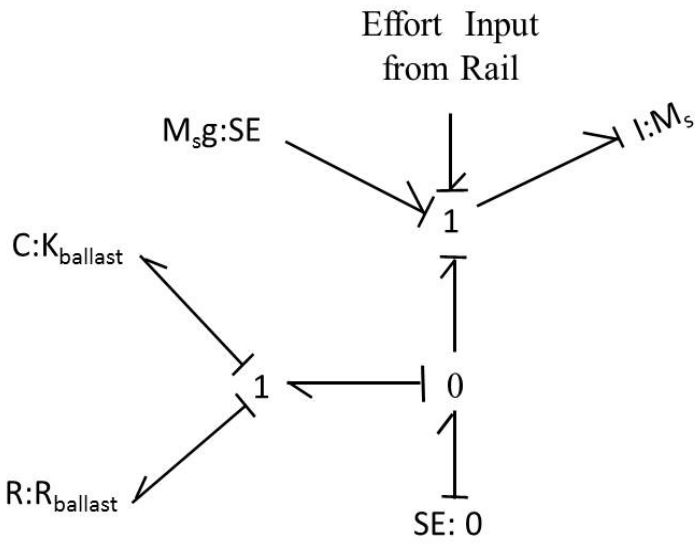


Fig. 4. Bond graph sub-model of sleeper and ballast (SB)

The coupled bond graph model of the coupled vehicle-track system is developed by the networking of the sub model of rail vehicle (RV), rail beam (RB) and sleeper and ballast (SB) and is shown in Fig. 5. In the integrated model total number of SB sub-model used is equal to no. Of sleepers (NS) ranging from SB_1, SB_2, ...SB_{NS}.

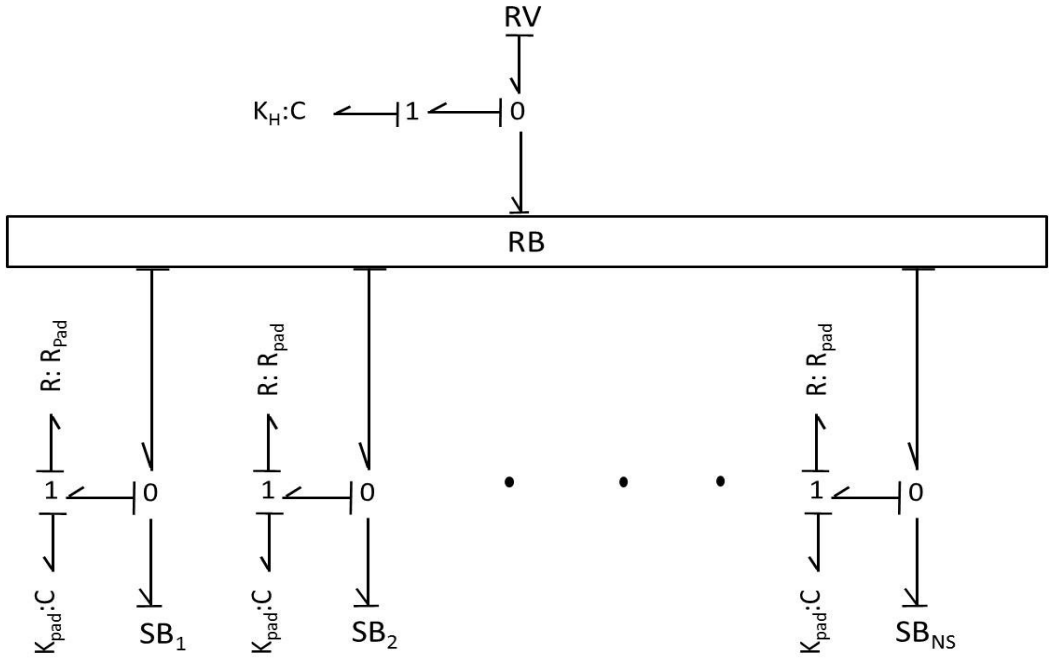


Fig. 5. Coupled bond graph representation of the coupled vehicle-track system

TABLE 1: Vehicle and track parameters [3, 22]

S.No.	Parameter	Nomenclature	Values
1	Rail flexural rigidity	EI	4.8645×10^6 Nm^2
2	Track Mass	M	119 kg/m
3	Foundation damping	R_f	2.76×10^4 Ns/m^2
4	Foundation stiffness	K_f	4.0×10^7 N/m^2
5	Wheel set mass	M_w	500 Kg
6	Static wheel-rail contact force	P_0	82000 N
7	Hertzian stiffness	K_H	1.0×10^{11} N/m$^{3/2}$
8	Wheel flat depth	D_f	2.15 mm
9	Wheel flat length	L_f	150 mm
10	Wheel speed	V	117 km/hr
11	Sleeper spacing	L_s	0.79 m
12	Rail mass in unit length	M_r	56 kg/m
13	Pad stiffness	K_{pad}	200×10^6 N/m^2
14	Pad damping*	R_{pad}	21.8×10^3 Ns/m^2
15	Ballast stiffness	$K_{ballast}$	31.6×10^6 N/m^2
16	Ballast damping*	$R_{ballast}$	21.8×10^3 Ns/m^2
17	Sleeper mass*	M_s	50 kg

**Parameters estimated based on parameters given by Newton and Clark [6]*

4. SIMULATION STUDY

The developed model is validated by comparing the contact force obtained in the current model and as reported by Newton and Clark [6] and Sadeghi et.al. [22](Fig. 6). The rail was assumed as a beam supported on an resilient base in these works and the vehicle as a wheel moving with speed of 117 Km/hr in these works So, to make the current model comparable with these models, SB sub model is removed from the coupled vehicle-track system model. The parameters obtained from the reported study [6, 22] are given in Table 1. The results display a considerable agreement when comparing the peak value of impact force and period of vibration obtained in the current model with the Newton and Clark and Sadeghi et al models.

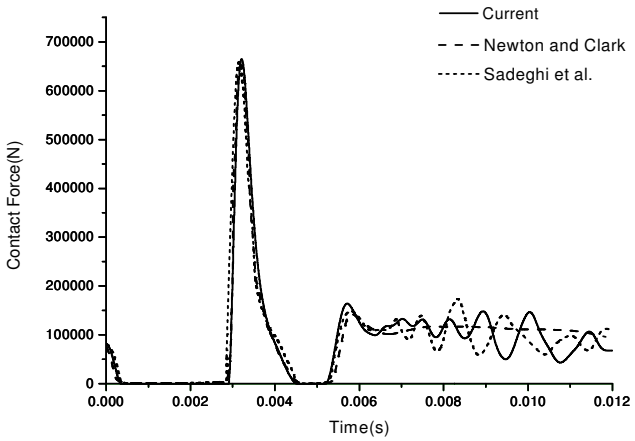


Fig. 6. Wheel/Rail contact force comparison for vehicle speed of 117 km/hr

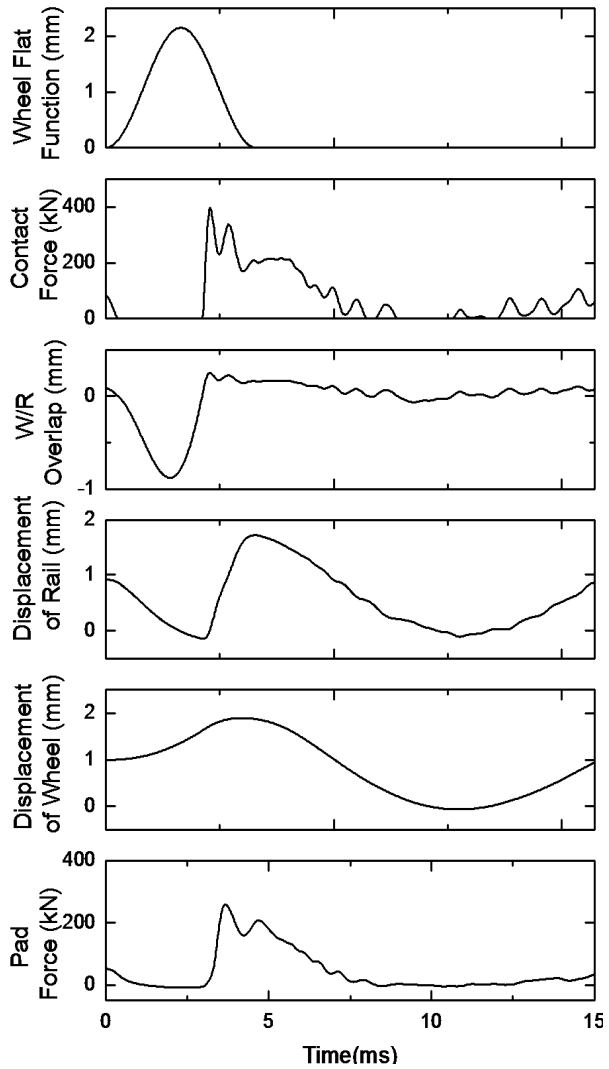
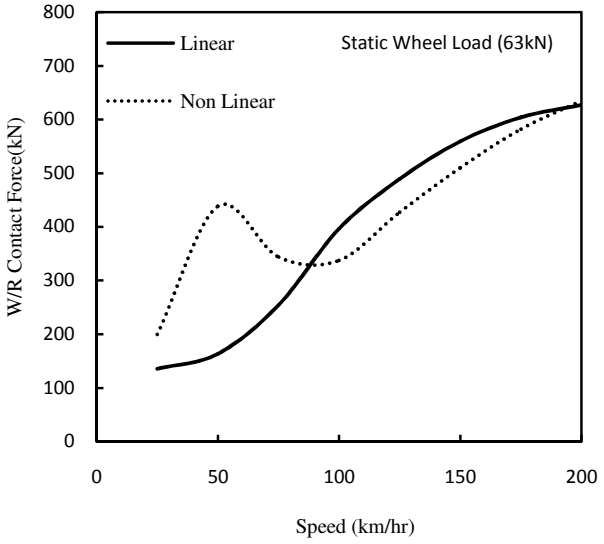
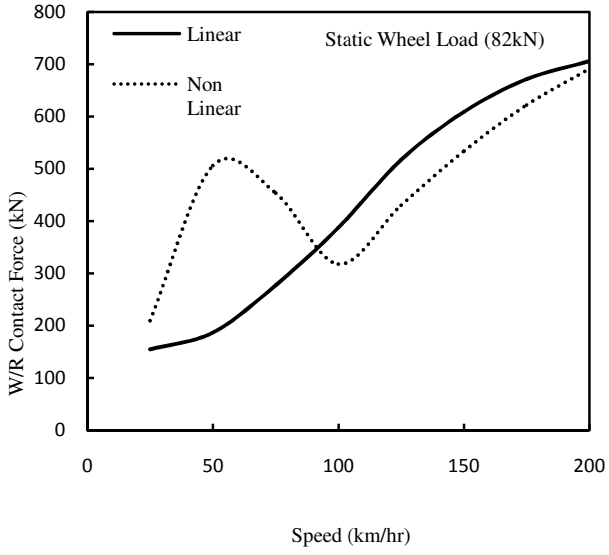


Fig. 7. Dynamic response of forces and displacement due to wheel flat for vehicle speed of 117 Km/hr.

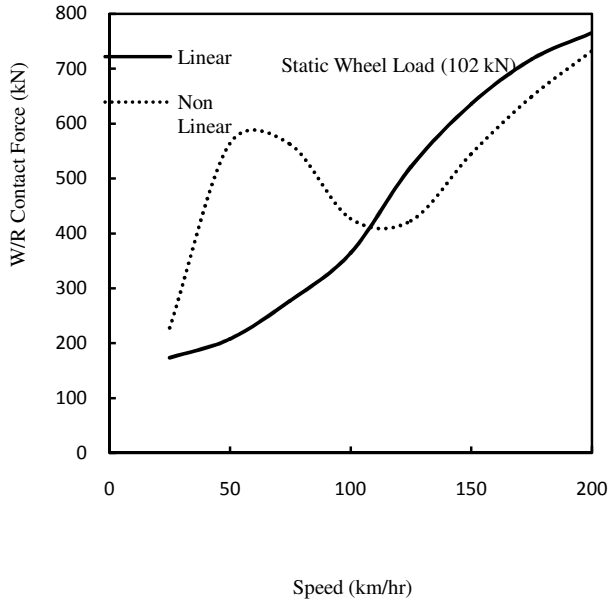
The presence of wheel flat defect causes significant wheel-rail impact loads, which may be many times higher than the static wheel load. This section presents the simulation results due to due to a single wheel flat of 150mm length and 2.15 mm deep. The wheel flat function along with the dynamic response of forces and displacements are presented in Figure 7. The plots revealed that after wheel flat appears the impact force gradually reduces to zero value. In such case, the wheel loses contact with the rail for a while and then again hits the rail with a large force when it strikes the rail.



(a)



(b)



(c)

Fig. 8. Effect of speed on peal contact force a) Static wheel load 63 kn; b) Static wheel load 82 kn; c) Static wheel load 102 kn

Figure 8 shows the response of impact for thetrack with linear and nonlinear pad and ballast at different speeds and static wheel load. It is evident from the plots that the peak impact forces rises with speed in case of thelinear track. However, for nonlinear track peak force increases up to 75 km/hr, then decreases with running speed up to 125 km/hr after which it again increases. The plots also revealed that peak value of impact forces from linear track model is considerably smaller up to 100 km/hr and beyond 100 km/hr the difference in the peak impact force isnot much. So, one can conclude that for wheel/trackinteraction problems assumption of thelinear track is inappropriate as the impact load is extremely large and the pad and ballast stiffness behaves in a nonlinear fashion.

5. CONCLUSIONS

In present paper, a coupled vehicle-track bond graph model has been developed to study the wheel/ track impact dynamics. Some nonlinear factors such as loss of wheel-rail contact, nonlinear behavior of pad and ballast is also considered. The response of contact force obtained from present model for beam on elastic foundation case has been compared with that reported by Newton and Clark [6] and Sadeghi et al[22] for similar studies. The impact force response shows considerably good agreement and establishes the correctness of the model.

The present model shows that; the vibration response of the linear track is normally lower than that of the nonlinear track as the predicted displacement is larger with a nonlinear track. It is observed that for wheel/track interaction problems assumption of the linear track is inappropriate as the impact load is extremely large and the pad and ballast stiffness being nonlinear changes considerably with the load. The current results have led to provide a better understanding of the interaction dynamics of the wheel/track interface dynamics with particular attention to the nonlinear behavior of pad and ballast.

REFERENCES

[1] AA Shabana, KE Zaazaa, H Sugiyama, (2007), Railroad vehicle dynamics: a computational approach: CRC press.

[2] R Dong, S Sankar, R Dukkipati, A finite element model of railway track and its application to the wheel flat problem, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 208(1994)61-72.

[3] Adkerr, Continuously supported beams and plates subjected to moving loads- a survey, SM Archives 6(1981)401-449.

[4] L Fryba, Dynamic interaction of vehicles with tracks and roads, Vehicle System Dynamics16(1987)129-138.

[5] K Knothe, SL Grassie, Modeling of railway track and vehicle/track interaction at high frequencies, Vehicle System Dynamics 22 (1993) 209 – 262.

[6] SG Newton, RA Clark, An investigation into the dynamic effects on the track ofwheel flats on railway vehicles" Journal of Mechanical Engineering Science 21 (1979) 287–297.

[7] M Fermer, JCO Nielsen, Vertical interaction between train and track with soft and stiff railpads—full-scale experiments and theory. Proc. Inst. Mech. Eng.209 (1995) 39–47.

[8] A Johansson, JCO Nielsen, Out-of-round railway wheels, wheel–rail contact forces and track response derived from the field tests and numerical simulation. Proc. Imeche, Part F: J. Rail and Rapid Transit 217 (2003) 135–146.

[9] YQ Sun, M Dhanasekar, A dynamic model for the vertical interaction of the railtrack and wagon system, International Journal of Solids and Structures 39 (2002) 1337–1359.

[10] RUA Uzzal, W Ahmed, S Rakheja, Analysis of pitch plane railway vehicle–track interactions due to single and multiple wheel flats, Proc. Imeche, Part F: J. Rail and Rapid Transit 223 (2009) 375-390.

[11] W Zhai, Z Cai, Dynamic interaction between a lumped mass vehicle and a discretely supported continuous rail track. Computers and Structures 63 (1997) 987-997.

[12] T Dahlberg, Dynamic interaction between train and nonlinear railway track model. In Proceedings of the 5th European Conference on Structural Dynamics, Munich, Germany, (2002) 1155-1160.

[13] TX Wu, DJ Thompson, The effects of track non-linearity on wheel/rail impact, Proc. Inst. Mech. Eng., Part F: J. Rail and Rapid Transit.218 (2003) 1-15.

[14] V Kumar, vrastogi, Investigation of vertical dynamic behaviour and modelling of a typical indian rail road vehicle through bond graph. World Journal of Modelling and Simulation 5 (2009) 130–138.

[15] D Karnopp, Bond graph for vehicle dynamics, Vehicle System Dynamics 1976; 5(3): 171-184.

[16] N Banerjee, R Karmakar, Bond Graph Modeling of Rail Wheelset on Curved Track, Simulation 83 (2007) 695-706.

[17] V Kumar, vrastogi, PM Pathak, Computational Analysis of Railway Vehicle on Irregular Tracks. In: Proceedings of The International Conference On Bond Graph Modeling And Simulation, Monterey, California, (2014) 34-40.

[18] D Margolis, T Shim, A bond graph model incorporating sensors, actuators, and vehicle dynamics for developing controllers for vehicle safety. Journal of the Franklin Institute 338 (2001) 21–34.

[19] R Merzouki, AK Samantaray, PM Pathak, Ould Bouamama, B. Intelligent Mechatronic Systems*:* Modeling, Control and Diagnosis. London, Springer, 2013.

[20] V Kumar, vrastogi, PM Pathak, Simulation for whole-body vibration to assess ride comfort of a low–medium speed railway vehicle, Simulation 93 (2017) 225-236.

[21] D Karnopp, DL Margolis, RC Rosenberg, System Dynamics: Modeling and simulation of mechatronic systems, 4th edition, Wiley, USA, 2006.

[22] Sadeghi, A. Khajehdezfuly, M. Esmaeili, D. Poorveis, Dynamic interaction of vehicle and discontinuous slab track considering nonlinear Hertz contact model, ASCE Journal of Transportation Engineering04016011 (2016)1–11

View publication stats

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/ACCESS.2018.2830651, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Effective Features to Classify Big Data using Social Internet of Things

Lakshmanaprabu SK[1], Shankar K[2], Ashish Khanna[3], Deepak Gupta[3], Joel J. P. C. Rodrigues[4], Senior Member, IEEE, Plácido R. Pinheiro[5], Member, IEEE, Victor Hugo C. de Albuquerque[5], Member, IEEE

[1] Department of Electronics and Instrumentation Engineering, B SAbdurRahman University, Chennai, India. Email: prabusk.leo@gmail.com
[2] Department of Computer science and Information Technology, Kalasalingam University, KrishnanKoil, India. Email: shankarcrypto@gmail.com
[3] Maharaja Agrasen Institute of Technology, GGSIP University, India. Email: deepakgupta@mait.ac.in, ashishkhanna@mait.ac.in
[4] National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí/MG, Brazil, Instituto de Telecomunicações, Portugal, ITMO University, St. Petersburg, Russia, University of Fortaleza (UNIFOR), Fortaleza/CE, Brazil. Email: joeljr@ieee.org
[5]Graduate Program in Applied Informatics, University of Fortaleza, Fortaleza/CE, Brazil. Email: placido@unifor.br; victor.albuquerque@unifor.br
Corresponding author: Joel J. P. C. Rodrigues (e-mail: joeljr@ieee.org).

ACKNOWLEDGEMENT This work was supported by the National Funding from the FCT – *Fundação para a Ciência e a Tecnologia* through the UID/EEA/50008/2013 Project, by the Government of Russian Federation, Grant 074-U01, by the Brazilian National Council for Research and Development (CNPq) via Grant No. 301928/2014-2, 304315/2017-6, and 309335/2017-5, and by Finep, with resources from Funttel, Grant No. 01.14.0231.00, under the Centro de Referência em Radiocomunicações - CRR project of the Instituto Nacional de Telecomunicações (Inatel), Brazil.

ABSTRACT: Social Internet of Things (SIoT) supports many novel applications and networking services for the IoT in more powerful and productive way. In the present work, we have introduced a hierarchical framework for feature extraction in SIoT big data using map-reduce framework along with supervised classifier model. Moreover, Gabor filter is used to reduce noise and unwanted data from the database, Hadoop Map Reduce have been used for mapping and reducing big databases, to improve the efficiency of the proposed work. Furthermore, the feature selection has been performed on filtered dataset by using Elephant Herd Optimization. The proposed system architecture has been implemented using Linear Kernel Support Vector Machine-based classifier to classify the data and for predicting the efficiency of the proposed work. The result of the proposed SIOT big data has been compared with the existing literature.

INDEX TERMS Internet of Things, Social Internet of Thing, Machine Learning, Big Data, Feature Selection.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

1. INTRODUCTION

Developing IoT enabled technologies and their solutions is a major challenge. However, IoT is about the pervasive collection and sharing of data towards a common goal [1]. In IoT, the data refers to attribute values such as variables or integer values; and event refers to when certain conditions are met or when certain states are reached [2, 4]. IoT service allows certain functions to be carried out through a predefined interface [5, 6]. Some researchers are particularly interested in identifying risk issues arising during discovering and integrating data within IoT [7]. Recently, the Social Internet of Things (SIoT) have been developed; another utilization of Internet of Things (IoT). The SIoT is a larger social network, connecting people and people, people and objects, and objects and objects [8]. Thus creating the opportunities, which brings several challenges to data processing systems for improving data collection, cleaning and storage, and performing real-time analytics [9]. Moreover, in the current scenario of Big Data, various standards and platforms have been introduced by relational database vendors and can be used for data aggregation as well as data analysis [10, 11]. The "big data and SIoT is the perfect representation of social systems and the IoT to characterize human progression [12]. Various feature selection algorithms have been proposed; which can be classified into two broad categories, i.e., filter approaches, and wrapper approaches. In filter-based approach, the filtration process is performed before classification process because of the independent usage nature of classification algorithms [13]. Linear Kernel Support Vector Machine-based classifier is used to classify the data. LK-SVM has numerous benefits like resolving little specimen; nonlinear and high dimensional pattern recognition [14]. Subsequently, the advanced things can't specifically apply to SIoT. In addition, heterogeneous nature of SIoT, different things communicating with each other under its umbrella devours a high volume of memory, power processing, along with high bandwidth [15].

The present work focuses on SIoT structure to ensure system safety with big data based on data classification model. To improve the performance of the system; an optimized feature extraction algorithm have been used. The data is classified by using LK-SVM classifier model. This work will be useful for the SIoT systems and will help to provide the guidance to the researchers of SIoT and big data.

The rest of the paper described as follows, section 2 performs state of the art in SIoT, classification and big data models and section 3 presents the detailed proposed work with different layers function includes block diagram. Section 4 comprises of simulation results of proposed work and performs comparison with the existing works and finally concludes our research work with future enhancement.

2. BRIEF LITERATURE REVIEW

Social Internet of Things (SIoTs) alludes to the quickly developing system of associated protests and individuals that can gather and exchange information using embedded sensors proposed by Hasan and Al-Turjman [16]. The researchers proposed a bio-inspired particle multi-swarm enhancement (PMSO) directing calculation to build, recoup and select disjoint ways that endure the disappointment while fulfilling Quality of Service (QoS) parameters. Multi-swarm technique empowers deciding the optimal directions in choosing the multipath directing while at the same time trading messages from all positions in the system. Results demonstrated that the procedure utilizing the qualities of all individual best data is a substantial technique for the motivations behind enhancing the PMSO execution.

Hussein et al. [17] have examined a novel administration structure in view of a reasoning approach for dynamic SIoT administrations revelation in smart spaces. Thus, reasoning about clients' situational needs, inclinations, and other social angles along with clients' encompassing condition were proposed for creating a file of situation-aware services which coordinates clients' needs. This reasoning approach is then actualized as a proof-of-idea model, specifically Airport Dynamic Social, inside a brilliant airport.

Nguyen et al. [18] investigated the issue in detail and propose a model-driven way to optimize an IoT application in regards to its non-practical prerequisites. A source code transformation, that updates the source code with the produced movable parameter values and then executes the compiler to make another binary image of the application. The investigational results shows that non-useful prerequisites, for example, power consumption and reliability can be enhanced generously at the time of optimization.

Multiple birth support vector machine is a novel machine learning calculation for multi-class characterization, which is considered as an augmentation of twin support vector machine by Zhang et al. [19]. To foresee another specimen initially decides an interim for each class depending on the separations between training tests and their hyperplanes and after that groups the new example relying upon the separations amongst hyperplanes and the new example which are in the relating interims. What's more, smoothing procedure is connected to the model, the first occasion when it was utilized as a part of multi-class twin support vector machine. The results on false datasets and UCI datasets demonstrated that the proposed algorithm was effective and has great order execution.

High-Performance Computing (HPC) solution has turned into a key issue and has pulled into consideration as of late been proposed by Ahmad et al. [20]. It presented a framework engineering that chooses Artificial Bee Colony (ABC). In addition, a Kalman filter was used as a part of Hadoop biological system that is utilized for evacuation of noise. A total four-level engineering was recommended that productively totals the information by removing superfluous information, and investigate the information using the proposed Hadoop-based ABC algorithm. To check the

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

effectiveness of the proposed calculations in the proposed framework engineering, we have executed the proposed framework utilizing Hadoop and MapReduce with the ABC calculation. ABC calculation was used to choose highlights, though; MapReduce was supported by a parallel algorithm that proficiently forms an enormous volume of data.

The Internet of Things (IoT) is over populated by a huge number of objects along with millions of interactions and services [21]. Mardini et al. [21] have recommended the SIoT; where each question in the IoT can utilize its companions' or friends-of-friends' connections to look for a particular administration. This is generally a moderate procedure on the grounds that every hub (object) is required to deal with a substantial number of companions. This work tends to the issue of link determination of companions and scrutinizes five techniques in the literature. It has been proposed a link determination procedure utilizing the Genetic Algorithm (GA) to locate the close optimal solution. The outcomes demonstrated a change in the inspected methodologies regarding a few parameters.

Thus, the main challenges are:

- Internet of Things gadgets, to interface with each other in a proficient way, consequently creating Big Data not deliver better reduction.
- The SIoT Big Data, to decrease the vitality utilization engaged with the SIoT at the time of communicating Data over the Internet.
- Advanced data analysis and intelligent approaches will provide utility to the SIoT with optimization and classifier model. [20]
- Develop a bio-inspired PSO routing algorithm achieve fast recovery from path failure by attempting to extend an existing approach. Optimization is required, in order to meet non-functional requirements in the application design phase and the problem of IoT is overcome by the proposed method [17, 18].

3. PROPOSED APPROACH

3.1 SIoT Data Collection

The developed things in the SIoT can able to detect the physical conditions, collects data, exchange or scatter information, process information for fitting applications, and communicate with different things. Henceforth, SIoT has been introduced with power innovation that aids in understanding the physical world and external jolts. While IoT can be considered of having ordinarily correspondence to the physical world by detecting or activating through huge numbers of various gadgets, SIoT worldview, this specimen appeared in figure 1, raises imperative issues regarding why and how to use these services as well as application.

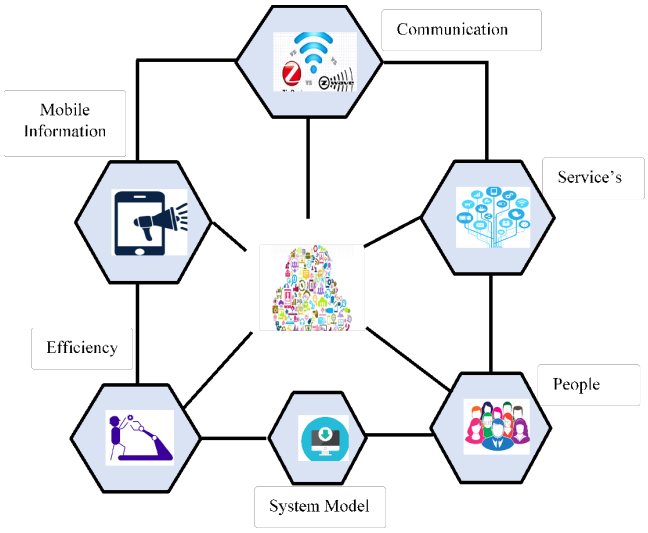


FIGURE 1. Sample Social IoT.

3.2 Filtering

Filtering is the key to the proposed approach, as for a particular dataset, some filter will be used to select the subset of dataset and, subsequently, the execution will remove noise and undesirable data.

GABOR Filter

Gabor is Gaussian function, which have been revised by complex sinusoids. Under specific conditions, the period of the reaction of Gabor channels is roughly linear. Gabor wavelet symbolizes a class of functions. The utilization of the Gabor channel bank can viably diminish the calculation, decrease the measurement and remove noisy information. A Gabor base is a Gaussian function, in which the exponential is adjusted with respect to the result of a Gaussian and an exponential. The two-dimensional Gabor function can be described using equation 1 and 2.

$$u(a,b) = g(a,b) * \exp^{-2\pi j f_y a} \tag{1}$$

Here

$$g(a,b) = \frac{1}{2\pi\sigma_a\sigma_b} * \exp^{\frac{-1}{2}\left(\frac{a^2}{\sigma_a^2}+\frac{b^2}{\sigma_b^2}\right)} \tag{2}$$

Gabor capacities are bandpass filters, which are Gaussians, focused on frequency in the spatial-area. Gabor channel with various frequencies and with introductions in various ways have been used to localize and extricate text area from data by removing noise.

3.3 Database Reduction

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

Hadoop MapReduce (HRM) comprises of a Job Tracker and a few Task Trackers. When a MapReduce task is executed, the Job Tracker divides it into smaller tasks (map as well as reduce), which can be processed by the task trackers and when the input dataset is given to a MapReduce data, it is split into autonomous data chunks using map task. The output of map task is the input of reduce task. The map function executes the technique of allocating each sample to the nearest focus while reduce task executes the new centers updating. As a result, the given data is mapped in a few groups and thus, the MapReducer reduces SIoT database using a threshold value.

MAPPING PROCESS:

Once the filtered data is converted into a set of data, a mapper can evaluate the nearest center point for every data. Maps are the individual tasks that change input data into the middle of the road information; such that the transformed intermediate data shouldn't have similar type as input data.

REDUCES PROCESS:

This process is known as diminish mission, which inputs the productivity from a map and unites the data tuples into a slighter group of tuples and results into a reduced set of values. Now, the reduce process combines the data created during the mapping process.

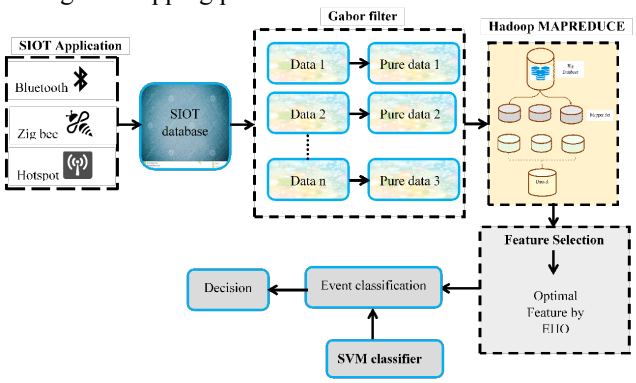


FIGURE 2. Proposed EH based classifying system for SIoT.

3.4 Feature Selection

Now, we have used feature selection technique to select optimal subset of features out of all the features. Here, we have used Elephant Herding Optimization (EHO) for feature selection and optimization. Elephant Herding Optimization searches the element space covetously and is constantly able to determine a better feature subset to optimize the given input data. It is used to improve classification.

Elephant Herding Optimization

In nature, the elephant is also considered as a social animal and the herd consists of several clans of female elephants and their calves. Under the influence of leader elephant or matriarch, the movement of every clan is defined. The Female Elephant (FE) use to live with their family groups whereas the Male Elephant (ME) separated when they grow up and live in contact with their family group utilizing low-frequency vibrations. Following assumptions are considered for herding:

- The population of elephants is divided into clans; each clan contains a fixed number of elephants.
- A fixed number of ME leaves their clan and live alone.
- Each clan moves under the leadership of a matriarch.

The group of matriarch holds the best solution in the herd of elephants while the worse solution is decoded from the position of the group of male elephants. Updating procedure of EHO has been shown below.

Elephant Position

In this step, the position of each elephant in different clans can be determined except the matriarch and male elephant that holds the best and worse solution respectively. For each clan C_i of P elephants. The position of i^{th} elephant $i = 1,2,....E$ and j^{th} clan $j = 1,2,....I$ is represented by $G_{i,j}$.

Movement update of fittest elephant of each clan

Elephants that move far from the group can be used to model the investigation. Every clan have some number of elephants with the most noticeably bad estimations of the target work has been moved to the new position.

Separating worst elephant's in clan

The worst elephant or male elephant will be separated from their family groups. During clan separation operation, the bits can be changed randomly since multiplication is performed with the random number. If the probability of randomness is given, then the number of bits to be changed can be estimated.

EHO Algorithm

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

(i) Initialization process
Choose feature attributes from the reduced database, EH parameters are scale factor and random values.
(ii) Fitness evaluation
Optimal feature selection consider accuracy as fitness function
(iii) Position of new elephant

$$G_{new,l_{i,j}} = G_{c\,i,j} + \alpha(G_{best,l\,i,j} - G_{l\,i,j}) \times r \quad \text{.....(3)}$$

(iv) Movement update
The position update for best fit in the clan is given by

$$G_{new,l_{i,j}} = \beta \times G_{center,l_j} \quad G_{center,l_j} = \sum_{i=1}^{n} G_{l_{i,j}} / n_l \quad \text{And}$$

.....(4)
(v) Separate worst position
The worst position updated as

$$G_{worst,l\,i,j} = G_{\min} + (G_{\max} - G_{\min} + 1) \times r \quad \text{.....(5)}$$

(vi) Memorize the best probable results
(vii) Exit if
The optimal solution with maximum accuracy reached.
else
Iteration = Iteration + 1.

Where,

$G_{new\,l\,i,j} \rightarrow$ Updated position,

$G_{l\,i,j} \rightarrow$ Old position,

$G_{best\,l\,i,j} \rightarrow$ Position of best in the clan,

α *and* β And $r \in 0\,to\,1$, $n_l \rightarrow$ The total number of elephants in each clan,

$G_{worst,ci,j} \rightarrow$ Worst male elephants in the clan and

$G_{\max}$ *and* $G_{\min} \rightarrow$ Maximum and minimum allowable boundary limits for the clan elephants.

In a given scenario of EHO, each clan source is linked with a bit vector, where the vector depends on the overall number of features. All the features are then evaluated and the feature, which has its value '1', is in the subset of optimal features. These optimal features are fed into the classifier model to improve precision and performance rate of the proposed model.

3.5 Classification

For the data classification, the optimal feature attributes are given as an input to support vector machine (SVM). The SVM could perceive new data as members or non-members of the functional class based on the expression data. The classifier has two models one as training and another as testing as shown in figure 3.

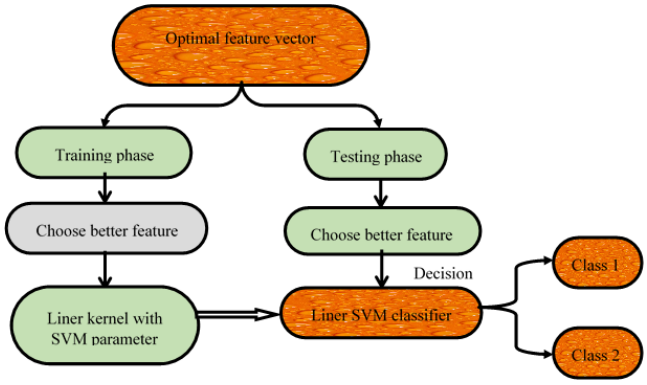


FIGURE 3. Linear kernel SVM-based classifier.

Training Model:

The scaling is essential for the variable (traits) with bigger variance. Training data can be used for separating hyperplane and if the training data are directly divisible, then choose these hyperplanes such that there are no focuses amongst them and to maximize their distance.

Linear Kernel SVM

A supposition is made that bigger the margin or distance between these parallel hyperplanes; the better the classifier error. Before considering N-dimensional hyperplanes, let us consider a simple 2-dimensional example. The linear kernel function appears beneath condition.

$$\sum_{i=1}^{N} \alpha_i . f_i . K(g, g_i) = 0 \quad (6)$$

Here g -vector drawn from the input space, assumed to be in the dimension

p_0, α_i is the Lagrange coefficient,

f_i is the equivalent target output and

$K(g, g_i)$ denotes the inner product of two vectors convinced in the feature space by the input vector.

$$K(g_i, g_j) = g_i^T g_j + C \quad (7)$$

Now, we have to apply the set of above parameter to the training dataset and get the classifier. An SVM maps the input into a high-dimensional feature space and finds a separating hyperplane.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

Testing Model:

Now, again apply the same set of parameters to the training dataset and get the classifier. Later on, use the classifier to classify the testing dataset to get the generalization accuracy.

4. RESULTS AND DISCUSSIONS

The proposed SIoT data classification with optimal feature selection implemented in Java programming language with JDK 1.7.0 in a windows machine with the configuration- 4 GB RAM and the operating system platform. The proposed model has been compared with the existing approaches like Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Genetic Algorithms (GA) based on the parameters like the total accuracy, the number of features, and classification results with performance metrics.

Database Description

The proposed optimal feature based classification models the training and testing process and considers the five different databases from the UCI machine learning repository. Each database contains some features/attributes and different classes; this information has been demonstrated in table 1. Each dataset have been analyzed as per the number of features. Description of the five databases used in the work are-

GPS Trajectories:

This dataset contains trip information regarding taxi stand, call origin details, unique taxi id, unique trip id and the Polyline. The Polyline is a string of GPS coordinates with each coordinate being obtained after every 15 seconds of the trip beginning with the pickup and ending with the drop-off.

Indoor User Movement Prediction from RSS:

This dataset represents a real-life benchmark in the area of Ambient Assisted Living applications. Input data contains temporal streams of radio signal strength (RSS) measured between the nodes of a WSN, comprising 5 sensors: 4 anchors deployed in the environment and 1 mote worn by the user.

Water Treatment Plant:

This dataset comes from the daily measures of sensors in an urban waste water treatment plant. The objective is to classify the operational state of the plant in order to predict faults through the state variables of the plant at each of the stages of the treatment process.

TABLE I
FIVE DATABASE ATTRIBUTE INFORMATION

Database	Number of Instances	Class label 1	Class Label 2
GPS Trajectories	163	Car	Bus
Indoor User Movement Prediction from RSS	13197	The location changing movements (1)	Location preserving movements (-1)
Water Treatment Plant	527	Healthy	Faulty
Hepatitis	155	Die	Live
Twitter Dataset for Arabic Sentiment Analysis	20000	Positive polarity	Negative polarity

Hepatitis:

The dataset contains 155 instances distributed between two classes die with 32 instances and live with 123 instances. There are 19 features or attributes, 13 attributes are binary while 6 attributes with 6-8 discrete values. The goal is to forecast the presence or absence of hepatitis virus.

Twitter Dataset for Arabic Sentiment Analysis:

The chosen tweets express some kind of emotions (positive or negative) and the goal of our model is to extract valuable information from such tweets with the end goal to determine the sentiment orientation of the inputted text. The months-long comment procedure of the tweets is physically led principally by two human specialists (local speakers of Arabic). On the off chance that the two specialists agree on the label of a specific tweet, then the tweet is doled out this label. Something else, a third master is counseled to break the tie.

Performance Metrics:

$$Sensitivity = \frac{TP}{TP + FN} \quad (8)$$

$$Specificity = \frac{TN}{TN + FP} \quad (9)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

Throughput: The throughput is the quantity of data (bits) conveyed within a time component (sec).

Energy Consumption (EC): The energy consumption is the standard of the energy devoted in data conveying from basic data to the target data within a time component.

TABLE 2
OPTIMAL FEATURES WITH ACCURACY FOR PROPOSED WORK

Database	Total Number of Features	Optimal Features	Accuracy
GPS Trajectories	15	6	78.21
Indoor User Movement Prediction from RSS	4	3	84.56
Water Treatment Plant	38	19	96.45
Hepatitis	19	7	92.1
Twitter Dataset for Arabic Sentiment Analysis	2	1	95.14

Five datasets such as GPS Trajectories, Indoor User Movement Prediction from RSS, Water Treatment Plant, Hepatitis and Twitter Dataset for Arabic Sentiment Analysis and their optimal set of selected features are detailed in the table 2. For example, water treatment plant has 38 number of features, among that, 19 optimal features are selected using the optimization algorithm with 96.45% accuracy.

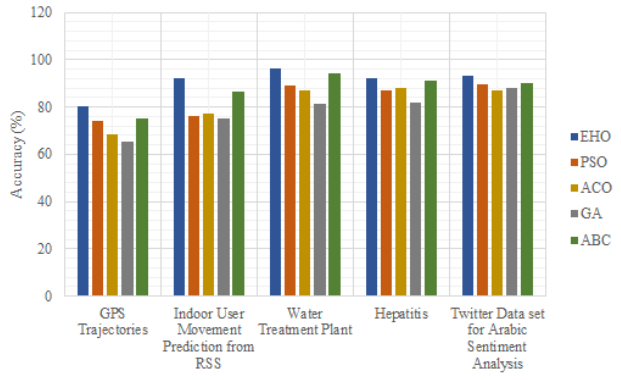


FIGURE 4. Comparative analysis for Feature selection

Figure 4 explains the comparative analysis of optimization algorithm in the feature selection process. For each database, the accuracy of optimal feature selection is compared with the algorithm like EHO, PSO, ACO, GA, and ABC. From the overall analysis, the proposed EHO based feature selection attains maximum accuracy in all the five databases compared to other related algorithms.

TABLE 3
COMPUTATIONAL TIME AND MEMORY FOR PROPOSED WORK

Data Size (Mb)	Time (Sec)	Memory (byte)
50	55	1209421
100	58	1324581
150	120	1425131
200	180	1485402
250	200	1528231
300	280	1651941

Table 3 describes the time required for the data computation and its memory allocation for different data size. For example, 100 Mb data size needs 58 seconds for data computation and its corresponding memory allocation is about 1324581 bytes. Similarly, for other data set, the computation time and memory size are revealed in the above table.

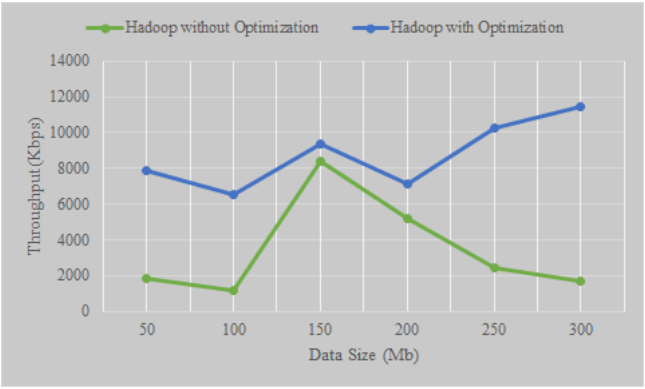


FIGURE 5. Throughput Vs Data Size

Figure 5 represents the throughput (Kbps) for various data size (Mb) such as 50, 100, 150, 200, 250 and 300. Throughput is analyzed with two conditions namely, Hadoop with optimization and Hadoop without optimization. On comparing both the analysis, Hadoop with optimization accomplishes maximum throughput in the range of 6000 to 12000.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

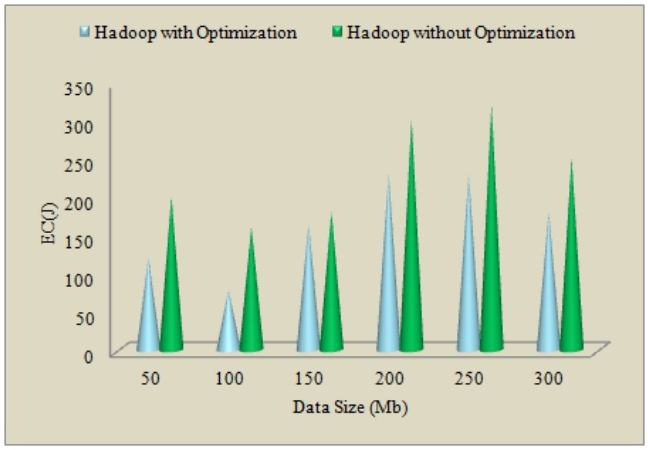


FIGURE 6. Energy Consumption Vs Data Size

The energy consumption (EC) of different data size is depicted in figure 6 under two conditions, namely Hadoop with optimization and Hadoop without optimization. The amount of energy consumed during the data convey from basic data to the target data is measured and compared. The graph clearly shows that the energy consumption is high in the Hadoop without optimization compared to Hadoop with optimization.

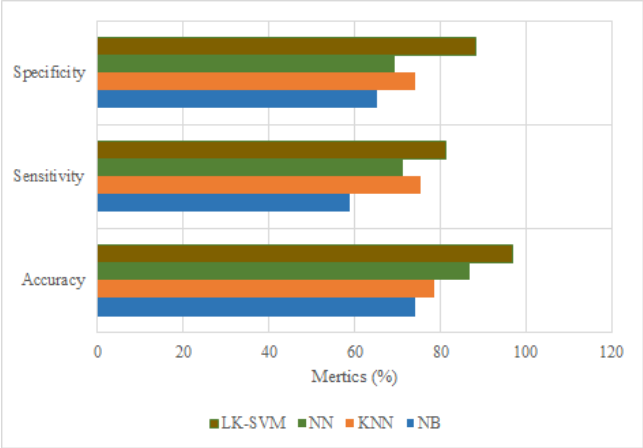


FIGURE 7. Performance metrics for various classifiers

Figure 7 illustrates the performance metrics of various classifiers such as LK-SVM, Neural Network (NN), K-Nearest Neighbour (KNN), and Naïve Bayes (NB). Compared to the existing classifiers, the proposed LK-SVM classifier attains more specificity and sensitivity. The SVM could classify the data as members or non-members of the class based on the expression data. The highest accuracy of data classification is 98.86%, which is achieved, in the proposed LK-SVM.

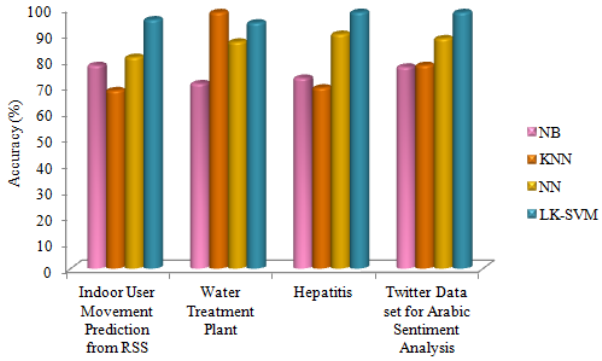


FIGURE 8. Comparative analysis for Classification Performance

Figure 8 represents the accuracy of the different classifier for the four different datasets i.e. Indoor User Movement Prediction from RSS, Water Treatment Plant, Hepatitis and Twitter Dataset for Arabic Sentiment Analysis. For the four datasets, the maximum accuracy is accomplished in the proposed LK-SVM compared to other classifiers.

5. CONCLUSION

The study scrutinized the classification of Social Internet of Things (SIoT) with big data map-reduce system. Here, Gabor filter has been used to reduce noise and the redundant data from the database. Furthermore to improve the efficiency of the proposed work, Hadoop MapReduce have been analyzed for mapping the data. The optimal features were selected from the database using the nature inspired algorithm for data classification. In this research work, the Elephant Herding Optimization (EHO) has been used to select the optimal subset of features from each dataset. Finally, Linear Kernel Support Vector Machine (LK-SVM) classifier was proposed to classify the data as members or non-members of the class based on the expression data. The result shows that the proposed LK-SVM with EHO attains a maximum accuracy of 98.86% compared to other existing approaches. The researcher can use the proposed work to improve the performance measures of big data classification based on SIoT by using advanced classifier with the hybrid optimization algorithms.

An application of the Internet of things in a social network that includes a smart device that can push its status messages on Twitter and in few areas like telecommunication, retail and finance have been early adopters of big data analytics.

REFERENCES

1. Guangjie Han, Lin Zhou, Hao Wang, Wenbo Zhang and Sammy Chan,"A Source Location Protection Protocol Based on Dynamic routing in WSNs for the Social Internet of Things", Journal of Future Generation Computer Systems,pp.1-23,2017.
2. Javier Lopez, Ruben Rios, FengBao and Guilin Wang,"Evolving privacy: From sensors to the Internet of Things", Journal of Future Generation Computer Systems, Vol. 75,pp.46–57, 2017.
3. Olayinka Johnny and SteliosSotiriadis,"Towards a social graph approach for modeling risks in Big Data and Internet of Things

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

(IoT)" In Intelligent Networking and Collaborative Systems (INCoS), 2014 International Conference on, pp. 439-444. IEEE, 2014.
4. Paul, Anand, Awais Ahmad, M. MazharRathore, and SohailJabbar. "Smartbuddy: defining human behaviors using big data analytics in thesocial internet of things." IEEE Wireless Communications,Vol.23, No.5,pp.68-74,2016.
5. Yim, Hyung-Jun, DongminSeo, Hanmin Jung, Moon-Ki Back, InA Kim, and Kyu-Chul Lee. "Description and classification for facilitating interoperability of heterogeneous data/events/services in the Internet of Things." Neurocomputing,pp.1-18, 2017.
6. Song, Zhiting, Yanming Sun, Jiafu Wan, Lingli Huang, Yan Xu, and Ching-Hsien Hsu. "Exploring robustness management of social internet of things for customization manufacturing." Future Generation Computer Systems,pp.1-11,2017.
7. Ochoa, Sergio F., Giancarlo Fortino, and Giuseppe Di Fatta. "Cyber-physical systems, internet of things and big data." Future Generation Computer Systems,Vol.75,pp. 82-84,2017.
8. Wasi-ur-Rahman, Md, NusratSharmin Islam, Xiaoyi Lu, Jithin Jose, HariSubramoni, Hao Wang, and Dhabaleswar K. DK Panda. "High-performance RDMA-based design of Hadoop MapReduce over InfiniBand." In Parallel and Distributed Processing Symposium Workshops &Ph.D. Forum (IPDPSW), 2013 IEEE 27th International, pp. 1908-1917. IEEE, 2013.
9. Ahmed, Ejaz, IbrarYaqoob, Ibrahim AbakerTargioHashem, Imran Khan, Abdelmuttlib Ibrahim Abdalla Ahmed, Muhammad Imran, and Athanasios V. Vasilakos. "The role of big data analytics in theInternet of Things." Computer Networks,pp.1-22,2017.
10. Rathore, M. Mazhar, Awais Ahmad, and Anand Paul. "IoT-based smart city development using big data analytical approach." In Automatica (ICA-ACCA), IEEE International Conference on, pp. 1-8. IEEE, 2016.
11. Guo, Kehua, Yayuan Tang, and Peiyun Zhang. "CSF: Crowdsourcing semantic fusion for heterogeneous media big data on the internet of things." Information Fusion,Vol.37,pp.77-85,2017.
12. FurqanAlama, Rashid Mehmood, IyadKatiba and AiiadAlbeshri,"Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT)", Procedia Computer Science,Vol.98,pp. 437-442,2016.
13. Huang, Chung-Ming, Chih-Hsiang Shao, Shou-ZhiXu, and Huan Zhou. "The Social Internet of Thing (S-IOT)-based Mobile Group Handoff Architecture and Schemes for Proximity Service." IEEE Transactions on Emerging Topics in Computing,pp.1-12,2017.
14. Jian, An, XiaolinGui, Wendong Zhang, and Xin He. "Social relation predictive model of mobile nodes in theinternet of things." ElektronikainElektrotechnika, Vol.19,No.481-86,2013.
15. Rasooli, A.,and Down, "A hybrid scheduling approach for scalable heterogeneous Hadoop systems". In High-Performance Computing, Networking, Storage,and Analysis (SCC), 2012 SC Companion, pp1284-1291.2012.
16. Mohammed ZakiHasan, and Fadi Al-Turjman,"SWARM-based data delivery in Social Internet of Things" Hasan, Mohammed Zaki, and Fadi Al-Turjman. "SWARM-based data delivery in Social Internet of Things." Future Generation Computer Systems,pp. 1-15,2007.
17. Hussein, Dina, Son N. Han, GyuMyoung Lee, Noel Crespi, and Emmanuel Bertin. "Towards a dynamic discovery of smart services in the social internet of things." Computers & Electrical Engineering,Vol.58,pp. 429-443,2017.
18. Nguyen, XuanThang, Huu Tam Tran, HarunBaraki, and Kurt Geihs. "Optimization of non-functional properties in theinternet of things applications." Journal of Network and Computer Applications,Vol.89,pp.120-129,2017.
19. Zhang, Xiekai, Shifei Ding, and Yu Xue. "An improved multiple birth support vector machine for pattern classification." Neurocomputing, Vol.225, pp.119-128,2017.
20. Ahmad, Awais, Murad Khan, Anand Paul, Sadia Din, M. MazharRathore, GwanggilJeon, and Gyu Sang Choi. "Toward modeling and optimization of features selection in Big Data based social Internet of Things." Future Generation Computer Systems,pp.1-16,2017.
21. Mardini, Wail, YaserKhamayseh, MuneerBaniYassein, and Montaha Hani Khatatbeh. "Mining Internet of Things for intelligent objects using agenetic algorithm." Computers & Electrical Engineering, pp.1-12,2017.



LAKSHMANAPRABU SK received his BE degree in Electronics and Instrumentation Engineering from Anna University, Chennai, India, in 2009. He completed his Post Graduate degree in Industrial Engineering in the year 2011 from Anna University, Trichy. He is currently pursuing his Ph.D. degree in Multivariable Optimal Control in the Department of Electronics and Instrumentation. His area of interest includes multivariable process control, fuzzy logic control, industrial automation, fractional order control system, model-based development, hardware in the loop test. He is currently working as asenior research fellow at B.S. AbdurRahman University, Chennai, India.



SHANKAR K is an assistant professor in the Department of Computer Science and Information Technology at the Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu, India. He received his master degrees ofMaster of Computer Applications, Master of Philosophy in Computer Science and aPh.D. degree in computer science from Alagappa University, Karaikudi, India. He has several years of experience working in the research, academia, and teaching. His current research interests include Cryptography and Network Security, Cloud security, Image Processing and Soft Computing Techniques.



ASHISH KHANNA is a highly qualified individual with around 15 years of rich expertise in Teaching, Entrepreneurship, and Research & Development with specialization in Computer Science Engineering Subjects. He received his Ph.D. degree from National Institute of Technology, Kurukshetra. He has completed his M. Tech. in 2009 and B. Tech. from GGSIPU, Delhi in 2004. Been a part of various seminars, paper presentations, research paper reviews, conferences as convener and session chair, guest editor in journals & have co-authored several books in publication house and papers in national, international journals and conferences. He has published many research papers in reputed journals and conferences. He also has papers in SCI-indexed and Springer journals. Guest editor in Springer Conference at ICDMAI-2018 and managing a special session on “Computational Intelligence for Data Science”. A reviewer in some SCI indexed Journals like “Cluster Computing” under Springer and IEEE conferences. Guest Editor in IEEE conference- IC3TSN-2017 and managing a special session on Parallel and Distributed Network-based Computing Systems.

Displayed vast success in continuously acquiring new knowledge and applying innovative pedagogies and have always aimed to be an effective educator and have a global outlook which is the need of today. He has co-authored 10 Text books and edited books i.e. Distributed Systems, Java Programming and Website Development, Java Programming, Computer Graphics, Computer Graphics and Multimedia, Computer Networks, Computer Networks and Data Communication Networks, Success Mantra for IT interviews, and Fundamental of Computing. He has also an edited book in Lambert publication. He is also a reviewer and session chair of IEEE international conference ICCCA 2016, 2017. He has designed the syllabus for cloud computing, Java Programming and Distributed Systems for GGSIPU, Delhi. He recently successfully managed Smart India Hackathon in 2017 at MAIT GGSIPU with teams under him winning prices at their

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

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/ACCESS.2018.2830651, IEEE Access

respective events. His research areas include Distributed Computing, Distributed Systems, Cloud Computing, Vehicular Ad hoc Networks and opportunistic networks.
Currently, he is associated with some Springer and IEEE conferences and managing special sessions for them and looking forward for some more challenging tasks.



DEEPAK GUPTA received his Ph.D from Dr APJ Abdul Kalam Technical University (AKTU) and Master of Engineering from Delhi College of Engineering, in 2017 and 2010, respectively. He is currently Assistant Professor at Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, GGSIP University, Delhi, India. He has published many scientific papers in SCI journals like JoCS, IAJIT, FGCS and etc. In addition he has authored/edited over 25 books (Elsevier, Springer, IOS Press, Katson) and 06 journal special issues. His research area includes Software Usability, Human-Computer Interaction, Nature-Inspired Computing, Machine Learning, and Soft Computing. He is convener and organizer of ICICC-2018 springer conference. He has also started a research unit under the banner of "Universal Innovator". He is also associated with various professional bodies like ISTE, IAENG, IACSIT, SCIEI, ICSES, UACEE, Internet Society, SMEI, IAOP, IAOIP etc.



Joel J.P.C. Rodrigues [S'01, M'06, SM'06] is a professor and senior researcher at the National Institute of Telecommunications (Inatel), Brazil and senior researcher at the Instituto de Telecomunicações, Portugal. He received the Academic Title of Aggregated Professor in informatics engineering from UBI, the Habilitation in computer science and engineering from the University of Haute Alsace, France, a PhD degree in informatics engineering and an MSc degree from the UBI, and a five-year BSc degree (licentiate) in informatics engineering from the University of Coimbra, Portugal. He is the leader of the Internet of Things Research Group (CNPq) and the NetGNA Research Group, the President of the scientific council at ParkUrbis – Covilhã Science and Technology Park, the Past-Chair of the IEEE ComSoc Technical Committee on eHealth, the Past-chair of the IEEE ComSoc Technical Committee on Communications Software, Steering Committee member of the IEEE Life Sciences Technical Community and Publications co-Chair, and Member Representative of the IEEE Communications Society on the IEEE Biometrics Council. He is the editor-in-chief of the International Journal on E-Health and Medical Communications, the editor-in-chief of the Recent Advances on Communications and Networking Technology, the editor-in-chief of the Journal of Multimedia Information Systems, and editorial board member of several high-reputed journals. He has authored or coauthored over 500 papers in refereed international journals and conferences, 3 books, and 2 patents. He had been awarded several Outstanding Leadership and Outstanding Service Awards by IEEE Communications Society and several best papers awards. Prof. Rodrigues is a licensed professional engineer (as senior member), member of the Internet Society, an IARIA fellow, and a senior member ACM and IEEE.

Victor Hugo C. de Albuquerque has a PhD in Mechanical Engineering with emphasis on Materials from the Federal University of Paraíba (UFPB, 2010), an MSc in Teleinformatics Engineering from the Federal University



of Ceará (UFC, 2007), and he graduated in Mechatronics Technology at the Federal Center of Technological Education of Ceará (CEFETCE, 2006). He is currently Assistant VI Professor of the Graduate Program in Applied Informatics, and coordinator of the Laboratory of Bioinformatics at the University of Fortaleza (UNIFOR). He has experience in Computer Systems, mainly in the research fields of: Applied Computing, Intelligent Systems, Visualization and Interaction, with specific interest in Pattern Recognition, Artificial Intelligence, Image Processing and Analysis, as well as Automation with respect to biological signal/image processing, image segmentation, biomedical circuits and human/brain-machine interaction, including Augmented and Virtual Reality Simulation Modeling for animals and humans. Additionally, he has research at the microstructural characterization field through the combination of non-destructive techniques with signal/image processing and analysis, and pattern recognition. Prof. Victor is the leader of Computational Methods in Bioinformatics Research Group. He is editorial board member of the IEEE Access, Computational Intelligence and Neuroscience, Journal of Nanomedicine and Nanotechnology Research, and Journal of Mechatronics Engineering, and he has been Lead Guest Editor of several high-reputed journals, and TPC member of many international conferences. He has authored or coauthored over 160 papers in refereed international journals, conferences, 4 book chapters, and 4 patents.



Placido Rogerio Pinheiro obtained the B.S. (1979) and Licentiate (1984) degrees in mathematics by the Federal University of Ceara, Ceara, Brazil, and in electrical engineering (1983) by the University of Fortaleza, Fortaleza, CE, Brazil, the M.S. (1990) degree in mathematics by the Federal University of Ceara, Ceara, Brazil, and Ph.D. (1998) degree in systems engineering and computing degree by the Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil. Currently, he is a Titular Professor of University of Fortaleza and Associate Professor of the State University of Ceara. He has experience in industrial processes modeling applying mathematical programming and multicriteria. His academic formation allows publishing in applied mathematics area, with an emphasis on discrete and combinatorics mathematics, working mainly in mathematical programming and multicriteria.

2169-3536 (c) 2018 IEEE. Translations and content mining are permitted for academic research only. Personal use is also permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934



Effective methods of waste management in India

R.S. Mishra, P.V Ram Kumar, Amit Pal, Mayank Kumar Singh

Department, of Mechanical & Production Engineering, Delhi Technological University Delhi-110042

Abstract

Municipal Solid Waste (MSW) is a major concern for all of the world including our country on this date, a lot of solid waste is being produced in India on a daily basis and this waste causes harm to environment and human beings too, so we need to have good methods to reduce the detrimental effect of these wastes to the environment. This paper has discussed some of the methods that are used for waste management in India namely Burying of solid wastes, Incineration of solid waste, The 'R' solution and Biological treatment of organic waste, paper also discusses the different challenges faced by these methods in India and how to minimize the waste produced. Solid waste management is one among the basic essential services provided by municipal authorities in the country to keep urban centers clean. However, it is among the most poorly rendered services in the basket. The systems applied are unscientific, outdated and inefficient; population coverage is low; and the poor are marginalized. Waste is littered all over leading to insanitary living conditions. Municipal laws governing the urban local bodies do not have adequate provisions to deal effectively with the ever- growing problem of solid waste management. With rapid urbanization, the situation is becoming critical. The urban population has grown fivefold in the last six decades with 285.35 million people living in urban areas as per the 2001 Census.
.

© 2017 ijrei.com. All rights reserved

Keywords: Waste Management, MSW, Land fill Management, Incineration

1. Introduction

In India, a proper waste management system is urgent necessary for the following reasons:

- To control different types of pollution, i.e., air pollution, soil pollution, water pollution etc.
- To stop the spread of infectious diseases.
- To conserve all our environmental resources, including forest, minerals water etc.
- To recycling of hazardous wastes for further production.

For this various waste management methods should be studied so that they can be used for minimising harm from the waste in our country

Although there are alarming predictions that global waste will nearly double to 2.2 billion annual tonnes by 2025 [1], and although almost all the cities in the world have problems to meet their waste reduction goals [2], urban waste has not get the required attention as much as water or energy in city planning are getting [3], nor has commensurate attention been given to the manifold infrastructures required to contend with this mounting problem and also waste consumes space, the problem of waste and its negative impacts have yet to be fully studied in planning and environmental design [4]. MSW, which refers to the materials disposed in urban areas which includes pre-dominantly household waste and commercial waste which is collected and disposed by municipalities, has long posed hostility to environmental quality and human health [5]. In past following years, rigorous growth of population, fast economic development and constant urbanization have led to a very large amount of wastes generation from industrial activities and household consumption, especially in capitals and major urban centres [6]. These urban wastes can be classified as municipal solid waste (MSW), industrial and domestic waste water and animal waste, etc [7]. Waste management is very important for sustainable development of a country and this paper have reviewed some of waste management techniques used in India.

Solid waste is any unwanted or discarded material we produce that is not a liquid or gas or any garbage, refuse, sludge from a waste treatment plant or air pollution control facility, and other discarded material (including solid, liquid, semi-solid or

Corresponding author: R.S.Mishra
Email Id: hod.mechanical.rsm@dtu.ac.in

contained gaseous material) generated from any industrial, commercial or community activities, mining or agricultural operations which can be categorized in two types i.e. Municipal Solid Waste and Industrial Solid Waste which is produced indirectly by industries that supply people with goods and services.

India is experiencing high urbanization, currently 31 percent as per Census 2011, contributing to 11 percent of the world population and having 53 metropolitan cities which may jump to 87 in 2031 [8]. It was estimated in 2014 by Planning Commission that urban India will generate 2,76,342 TPD by 2021, 4,50,132 TPD by 2031 and 11,95,000 TPD of MSW by 2050. (Planning Commission, 2014). The report showed that MSW then contained 51 percent of organic waste, 32 percent inert and non-organic waste collectively. Plastics, paper, and glass constitute 17 percent of waste which can be classified as recyclable wastes [9].

The improper disposal of municipal waste has a serious and dangerous impact on a wide range of areas. Garbage thrown in the street or in open spaces creates a public health hazard, while waste dumped near rivers, lakes and streams contaminates the water supply. Rubbish that is burned in the open rather than disposed of properly creates pollution and releases toxic fumes into the environment. Non-biodegradable materials thrown into open drains make their way into the sewerage system, clogging pipelines and damaging infrastructure. The hazards posed by the dumping of untreated hospital and industrial waste are even greater, with the release of pathogens and toxic compounds posing a grave threat not just to human life but also to plants and animals. Garbage dumped in the countryside is not simply an eyesore; entire landscapes are ruined and unique habitats for flora and fauna are lost.

All of these problems are common in India, where vast quantities of solid waste remain uncollected in the streets, along major roads, in empty plots of land, downhill slopes and in illegal dumps. Table 1 shows the composition of the MSW produced in India in different year.

The main obstacles in the way of growth in wide use of waste management methods are

- Lack of awareness among the population.
- Segregation of MSW is difficult in India.
- Lack of appropriate level funding.
- Lack of implementation of rules at ground level.
- Lack of coordination among Centre and State.

The report of the Task Force on Waste to Energy (Vol.I), 2014 shows that out of 29 only 22 states/UTs have sufficient infrastructure for processing and disposing waste and the rest of the States/UTs had made no progress till 2013. Of the 279 conventional composting, 138 vermi composting facilities, 172 biomethanation,29 RDF and 8 Waste to Energy plants which were established many are either closed or not performing on their potential [9].

Drawbacks in present solid waste management services

1.1 No Storage of Waste at Source

There is no practice of storing the waste at source in a scientifically segregated way. Citizens have not been educated to keep domestic, trade, and institutional bins for storage of waste at source and stop littering on the streets.

1.2 No System of Primary Collection from the Doorstep

There is no public system of primary collection from the source of waste generation. The waste discharged here and there is later collected by municipal sanitation workers through street sweeping, drain cleaning, etc. Street sweeping has, thus become the principal method of primary collection.

1.3 Irregular Street Sweeping

Even street sweeping is not carried out on a day-to-day basis in most cities and towns in India. Generally commercial roads and important streets are prioritized and rest of the streets are swept occasionally or not swept at all. Generally, no sweeping is done on Sundays and public holidays and a back log is created on the next working day.

The tools used for street sweeping are generally inefficient and out -dated. For instance, the broom with a short handle is still in use forcing sweepers to bend for hours resulting in fatigue and loss of productivity.

Traditional handcarts/tricycles are used for collection, which do not synchronize with the secondary storage systems. Waste is deposited on the ground necessitating multiple handling.

There are no uniform yardsticks adopted for street sweeping. Though, some states/cities have prescribed work-norms, these are not very scientific. Most of the cities allocate work to sanitation workers on ad hoc basis. The work distribution ranges between 200 metres to 1000 metres of street sweeping each day. Some sanitation workers are found under worked while some overburdened

.

1.4 Waste Storage Depots

As waste is collected through traditional handcarts/tricycles that can carry only a small quantity of waste at a time, there is a practice to set up depots for temporary storage of waste to facilitate transportation through motorized vehicles. Generally, open sites or round cement concrete bins, masonry bins or concrete structures are used for temporary bulk storage, which necessitates multiple handling of waste. Waste often spills over which is both unsightly as well as unhygienic.

1.5 Transportation of Waste

Transportation of waste from the waste storage depots to the disposal site is done through a variety of vehicles such as bullock carts, three-wheelers, tractors, and trucks. A few cities use modern hydraulic vehicles as well. Most of the transport vehicles are old and open. They are usually loaded manually.

Year	Biodegradables	Paper	Plastic/Rubber	Metal	Glass	Rags	Others	Inert
1996	42.21	3.63	0.60	0.49	0.60	-	-	45.13
2005	47.43	8.13	9.22	0.50	1.01	4.49	4.02	25.16
2011	42.51	9.63	10.11	0.63	0.96	-	-	17.00

The fleet is generally inadequate and utilization in optimal. Inefficient workshop facilities do not do much to support this old and rumbling squad of squalid vehicles. The traditional transportation system does not synchronize with the system of primary collection and secondary waste storage facilities and multiple manual handling of waste results.

1.6 Processing of Waste

Generally no processing of municipal solid waste is done in the country. Only a few cities have been practising decentralized or centralized composting on a limited scale using aerobic or anaerobic systems of composting. In some towns un-segregated waste is put into the pits and allowed to decay for more than six months and the semi-decomposed material is sold out as compost. In some large cities aerobic compost plants of 100 MT to 700 MT capacities are set up but they are functioning much below installed capacity. A few towns are practising vermi-composting on a limited scale.

1.7 Disposal of Waste

Disposal of waste is the most neglected area of SWM services and the current practices are grossly unscientific. Almost all municipal authorities deposit solid waste at a dump-yard situated within or outside the city haphazardly and do not bother to spread and cover the waste with inert material. These sites emanate foul smell and become breeding grounds for flies, rodent, and pests. Liquid seeping through the rotting organic waste called leachate pollutes underground water and poses a serious threat to health and environment.
Landfill sites also release landfill gas with 50 to 60 per cent methane by volume. Methane is 21 times more potent than carbon dioxide aggravating problems related to global warming. It is estimated by TERI that in 1997 India released about 7 million tons of methane into the atmosphere. This could increase to 39 million tons by 2047 if no efforts are made to reduce the emission through composting, recycling, etc

1.8 Waste

It is defined as "non-liquid, non-soluble materials ranging from municipal garbage to industrial wastes that contain complex & sometimes hazardous substances". Solid waste also include.

- Garbage
- Rubbish
- Demolition products
- Sewage treatment residue
- Dead animals
- Manure and other discarded material.

1.8.1 Types of Waste

1. Household waste as municipal waste.
2. Industrial waste as hazardous waste.
3. Biomedical waste or hospital waste as infectious waste.

1.8.2 Municipal Solid Waste

Municipal solid waste consist of

- household waste
- construction and demolition debris
- sanitation residue
- waste from streets

With rising urbanization and change in life style and food habits, the amount of municipal solid waste has been increasing rapidly and its composition changing.
Municipal waste is generated by households, commercial activities and other sources whose activities are similar to those of households and commercial enterprises.
Municipal waste is made up to residual waste, bulky waste, secondary materials from separate collection (e.g., paper and glass), household hazardous waste, street sweepings and litter collections. It is made up of materials such as paper, cardboard, metals, textiles, organics (food and garden waste) and wood. Municipal waste represents approximately 14% of all waste generated.

1.8.3 Hazardous Waste

Industrial and hospital waste is considered hazardous as they may contain toxic substances
Hazardous waste could be highly toxic to humans, animals and plants. They are

- corrosive
- highly inflammable or explosive

In the industrial sector the major generators of hazardous waste are the metal' chemical' paper, pesticide, dye and rubber goods industries. Direct exposure to chemicals in hazardous waste such as mercury and cyanide can be fatal.
Hazardous waste is a waste which may pose a substantial present or potential hazard to human health or the environment

when improperly treated, stored or disposed of, or otherwise mismanaged or cause or contribute to an increase in mortality, or an increase in irreversible or incapacitating illness.
Hazardous waste is typically the subject of special legislation and requires special management arrangements to ensure that hazardous waste is kept separate from and treated differently to non-hazardous waste.

1.8.4 Hospital Waste Or Biomedical Waste

Bio-medical waste means "Any waste which is generated during the diagnosis, treatment or immunization of human beings or animals or in research activities pertaining thereto or in the production or testing of biological"
It may includes wastes like sharp waste, pathological waste, pharmaceutical waste, genotoxic waste, chemical waste, and radioactive waste etc.

2. Waste Management Methods

2.1 Burying Solid Waste

Most of the world's MSW is buried in landfills that eventually are expected to leak toxic liquids into the soil and underlying aquifers.

Open dumps: are fields or holes in the ground where garbage is deposited and sometimes covered with soil. Mostly used in developing countries.

Sanitary landfills: method of controlled disposal of municipal solid waste (refuse) on land. The method was introduced in England in 1912 (where it is called controlled tipping). In this method the waste is first collected in thin layers that can be of 1m high and then it is compressed by heavy machines (e.g., bulldozers); many layers of these are then placed on top of one other and compressed again to form a refusal cell which can be up to 3 metres, thick. The compressed cell is then protected with a layer of compressed soil by putting it on top to act as a shield to avoid odours and windblown debris. Modern landfill sites are carefully covered and sealed with impermeable synthetic bottom liners to avoid pollution of groundwater and other environmental pollutions. When the landfill is filled to the top then a layer of clay or synthetic liner is placed over it to prevent water from entering. On top of it a final layer of soil is placed then compressed and graded. After sometimes some plants can be grown on it to increase its usefulness and maybe convert into garden or parks. Chakraborty M, Sharma C, Pandey J and Gupta P.K. Estimated energy potential of incineration process in Delhi's MSW, bomb calorimeter was used to determine calorific values of Delhi's MSW samples. The calorific values was found to be 0.058–0.078 kW/kg for segregated waste containing mixture of kitchen waste, litter with small branch of shrubs, etc. and 0.092–0.126 kW/kg for the bulk waste (excluding plastic, rubber, wooden materials and packing and other inert materials) [10].

Sanitary landfills are the ultimate means of disposal of all types of residual, residential, commercial and institutional waste as well as unutilized municipal solid waste from waste processing facilities and other types of inorganic waste and inerts that cannot be reused or recycled in the foreseeable future.Its main advantage is that it is the least cost option for waste disposal and has the potential for the recovery of landfill gas as a source of energy, with net environmental gains if organic wastes are land filled. The gas after necessary cleaning, can be utilized for power generation or as domestic fuel for direct thermal applications1. Highly skilled personnel are not required to operate a sanitary landfill.
Major limitation of this method is the costly transportation of MSW to far away landfill sites. Down gradient surface water can be polluted by surface run-off in the absence of proper drainage systems and groundwater aquifers may get contaminated by polluted leacheate in the absence of a proper leacheate collection and treatment system.

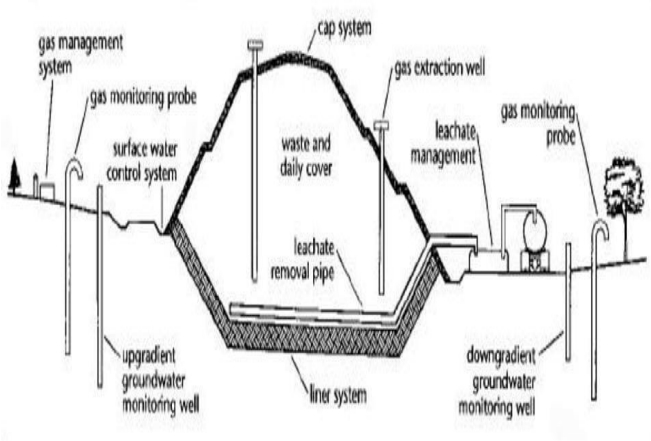


Fig.1- A typical Sanitary landfill [11].

3. Incineration or Burning of Solid waste

This method, commonly used in developed countries is most suitable for high calorific value waste with a large component of paper, plastic, packaging material, pathological wastes, etc. It can reduce waste volumes by over 90 per cent and convert waste to innocuous material, with energy recovery. The method is relatively hygienic, noiseless, and odourless, and land requirements are minimal. The plant can be located within city limits, reducing the cost of waste transportation.
This method, however, is least suitable for disposal of chlorinated waste and aqueous/high moisture content/low calorific value waste as supplementary fuel may be needed to sustain combustion, adversely affecting net energy recovery. The plant requires large capital and entails substantial operation and maintenance costs. Skilled personnel are required for plant operation and maintenance. Emission of particulates, SOx , NOx, chlorinated compounds in air and toxic metals in particulates concentrated in the ash have raised concerns. Combustion is a reaction which occurs between fuel and oxidant to produce heat. The fuel can be of three types

namely gaseous, liquid or solid. Upon combustion, chemical reactions between fuel and oxidant take place and finally the heat released from the reactions makes the process self sustaining [12]. Waste incineration is most appropriate for certain kind of waste which comprises of combustible MSW with non-biodegradable material and low moisture content [13]. In the world there are over 1000 large incinerators which converts waste to energy by using heat energy released from incinerator to boil water to make steam which can be used to produce electricity. Incineration is the process of direct burning of waste in the presence of excess air (oxygen) at temperatures of above 800˚C that liberates heat energy besides generating inert gases and ash. In practice, about 65–80% of the energy content of the organic matter can be recovered as heat energy, which can be utilized either for direct thermal applications, or for producing power via steam turbine generators [14]. This process can reduce the waste volume up to 90% [5] and the capital cost of this process is lower than the cost of biomethanation process.Fig.2 shows the construction and working of a typical incinerator. The main disadvantages of this technology is the high recurring cost of the tune of Rs. 5460/ton of waste processing in which 80% of this cost is spent on fuel to burn the waste [15]. To economically use this technology, the minimum calorific value of MSW required was found to be 0.07 kW/kg which is difficult to get due to presence of high moisture contents. Another major issue associated with this technology is the emission of gases like polychlorinated biphenyl (PCBs), hydrogen chloride, hydrogen fluoride, mercuric chloride, dioxins and furans (PCDD/F) which are extremely toxic and are considered to be carcinogens which are harmful and injurious for the animals and also to human health [16].

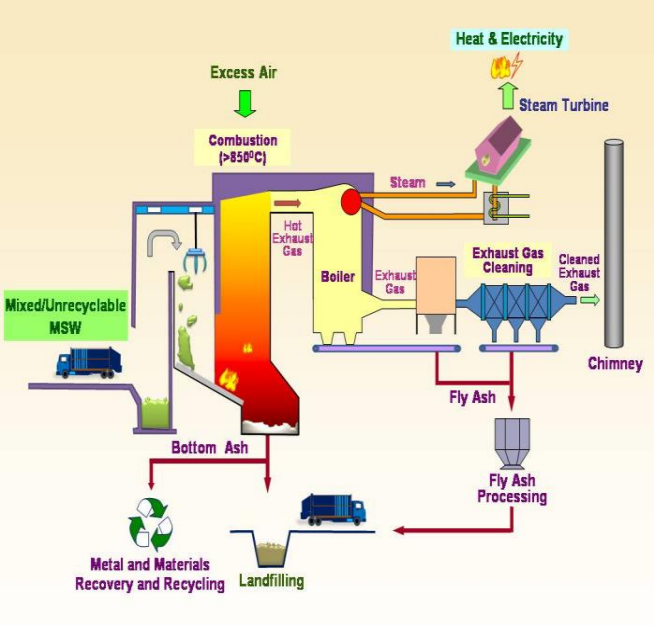


Fig.2- Waste-to-energy incinerator with pollution controls that burns mixed solid waste [17].

4. The R Solution

- *Refuse*: It means one should refuse to buy items that they don't need.
- *Reduce:* If refers to consume less and live a simpler and less stressful life by practicing simplicity.
- *Reuse*: use those items that can be used over and over.
- *Repurpose:* use the previously used item for another purpose instead of throwing it away.
- *Recycle*: paper, glass, cans, plastics and buy items made from recycled materials.

Reusing products is an important way to reduce resource use, waste, and pollution in developed countries. In developing countries where there are open dumps this can be dangerous as the scavengers are exposed to various toxins and harmful materials.

4.1 Primary (closed loop) recycling

In this type of recycling materials are turned into new products of the same type in other words primary recycling means that the recyclable material/product is recovered and reused without being changed in any way and usually for the very same purpose.

4.2 Secondary recycling

In this type of recycling materials are converted into different products. Used tires shredded and converted into rubberized road surface. Newspapers transformed into cellulose insulation. Composting biodegradable organic waste mimics nature by recycling plant nutrients to the soil. Recycling paper has a number of environmental (reduction in pollution and deforestation, less energy expenditure) and economic benefits and is easy to do. Many types of plastics are very hard to recycle due to their chemical composition and cost of appropriate recycling technique. Not very much material can be yielded from collecting individual plastic resin. The cost of new plastic resins in much less than recycled plastic resins which is due to low fossil fuel costs. There are new technologies in world that are making plastics biodegradable.

4.3 Tertiary recycling

Tertiary recycling can be defined as a process that involves chemical modification of the material used in order to use it again for a purpose.

5. Biological treatment of organic waste

5.1 Aerobic Composition

Roughly half of household waste is made up of food and garden waste. Most of this material could be composted to save landfill, improve soil condition and provide fertiliser in the garden at no cost. Composting is the method of breaking down

waste organic materials in a large container or heap. The decomposition occurs because of the naturally occurring micro-organisms, such as bacteria and fungi.

Aerobic composting is a process where adequate ventilation should be provided to allow respiration of microorganisms that release carbon dioxide into the atmosphere, therefore aeration of composting material is necessary for efficient decomposition. The oxygen saturation in the medium should not be lower than 5%, whereas 10% is the optimal level for efficient decomposition. Excessive aeration will cause a temperature drop and a great loss of moisture by evaporation, causing the decomposition process to stop. On the other hand, low aeration prevents enough water evaporation, generating excessive moisture and an anaerobic environment [18]. Most composters are designed to provide adequate aeration of the waste. In the event of insufficient aeration, it is necessary to stir the material.

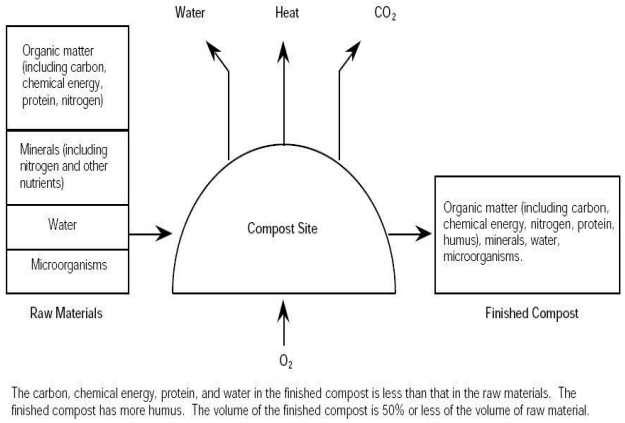


Figure 3. Aerobic Composting [20].

Composting is a technology known in India since times immemorial. Composting is the decomposition of organic matter by microorganism in warm, moist, aerobic and anaerobic environment. Farmers have been using compost made out of cow dung and other agro-waste. The compost made out of urban heterogeneous waste is found to be of higher nutrient value as compared to the compost made out of cow dung and agro-waste. Composting of MSW is, therefore, the most simple and cost effective technology for treating the organic fraction of MSW. Full-scale commercially viable composting technology is already demonstrated in India and is in use in several cities and towns. Its application to farm land, tea gardens, fruit orchards or its use as soil conditioner in parks, gardens, agricultural lands, etc., is however, limited on account of poor marketing.

Main advantages of composting include improvement in soil texture and augmenting of micronutrient deficiencies. It also increases moisture-holding capacity of the soil and helps in maintaining soil health. Moreover, it is an age-old established concept for recycling nutrients to the soil. It is simple and straightforward to adopt, for source separated MSW. It does not require large capital investment, compared to other waste treatment options. The technology is scale neutral. Composting is suitable for organic biodegradable fraction of MSW, yard (or garden) waste/waste containing high proportion of lignocelluloses materials, which do not readily degrade under anaerobic conditions, waste from slaughterhouse and dairy waste.

This method, however, is not very suitable for wastes that may be too wet and during heavy rains open compost plants have to be stopped. Land required for open compost plants is relatively large. Also, issues of methane emission, odour, and flies from badly managed open properly carried out there is possibility of toxic material entering the stream of MSW.

5.2 Vermi Composting

It is process in which earthworms are spread on semi-decomposed organic waste. Earthworms can eat upto five times of organic matter per day as compared to their body weight. Initially, biodegradable organic matter is decomposed through microbial enzymatic activity. 100 Million Tons/day capacity vermin-composting plant is located in Bangalore which is India's largest, while there are smaller plants in Hyderabad, Bangalore, Mumbai, and Faridabad.

Vermi-compost is the natural organic manure produced from the excreta of earthworms fed on scientifically semi-decomposed organic waste. A few vermi composting plants generally of small size have been set up in some cities and towns in India, the largest plant being in Bangalore of about 100 MT/day capacity. Normally, vermi-composting is preferred to microbial composting in small towns as it requires less mechanization and it is easy to operate. It is, however, to be ensured that toxic material does not enter the chain which if present could kill the earthworms.

5.3 Anaerobic digestion

Anaerobic decomposition of waste is also known as biomethanation process. It is one of the important and sustainable techniques for treatment of the biodegradable part of MSW in subtropical climates. In this process, stabilization occurs and biogas is liberated by the conversion of organic matter, which in turn can be used as energy. The biogas has 55–60% methane and it can be used as fuel for power generation. Government of India encourages biomethanation technology by utilizing industrial, agricultural and municipal wastes. Biomethanation has advantage due to low capital and operating costs compared to other methods but the main problem with this method is of amount space it uses for, as it requires the covering of waste for developing anaerobic conditions and that cover cannot be opened for next few years thereby making that space unavailable for further dumping of MSW and after 4–7 years, the methanation process becomes unviable due to low carbon content in the remaining waste. Therefore, very large space is required to operate biomethanation plants for treatment of city's daily MSW generation [10].

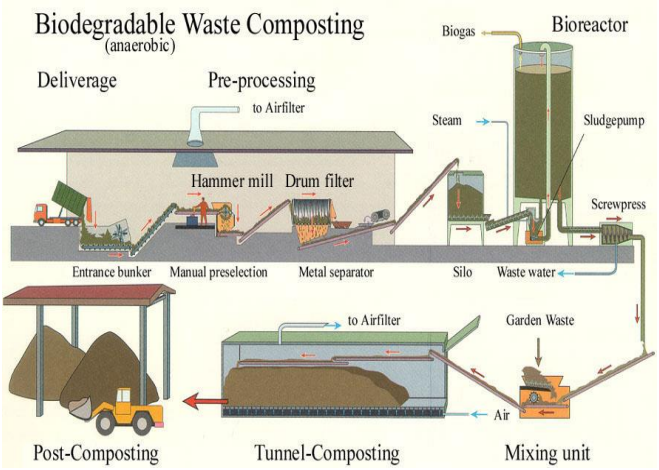


Figure 4. Anaerobic Composting [21].

Biomethanation is a comparatively well-established technology for disinfections, deodorization and stabilization of sewage sludge, farmyard manures, animal slurries, and industrial sludge. Its application to the organic fraction of MSW is more recent and less extensive. It leads to bio-gas/power generation in addition to production of compost (residual sludge). This method provides a value addition to the aerobic (composting) process and also offers certain other clear advantages over composting in terms of energy. This method is suitable for kitchen wastes and, other putrescible wastes, which may be too wet and lacking in structure for aerobic composting. It is a net energy-producing process (100–150 kWh per tonne of waste input). A totally enclosed system enables all the gas produced to be collected for use. A modular construction of plant and closed treatment needs less land area. This plant is free from bad odour, rodent and fly menace, visible pollution, and social resistance. It has potential for co-disposal with other organic waste streams from agro-based industry. The plant can be scaled up depending on the availability of the waste.

6. Refused derived fuel (RDF)

This is one of WTE process where installation costs as well as maintenance costs are lowest among all the available methods. Other benefits associated with this technology are the higher heating value (0.145–0.194 kW/kg) which also remains fairly constant in the presence of 3–8% moisture content [19], the homogeneity of physio-chemical composition, the ease of storage, handling, and transportation, the lower pollutant emissions and a reduced excess air requirement during combustion.

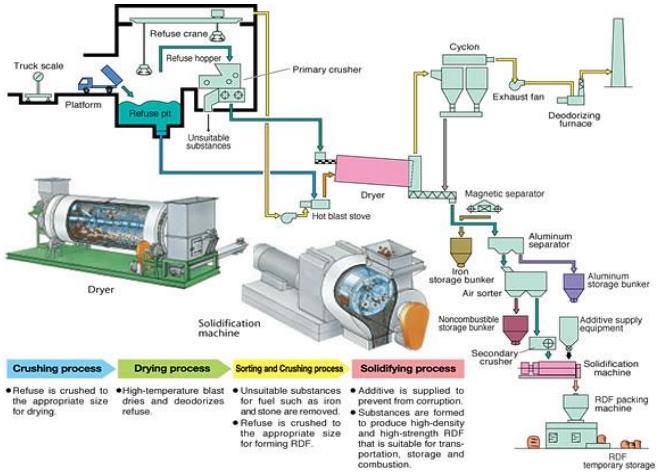


Figure 5. Refuse Derived Fuel Plant [22].

The high calorific value of RDF is due to its higher density (700 kg/m3) compared to the other forms of waste. It has been estimated that 160–180 ton of garbage are required to produce 40 ton of pellets [19]. Chakraborty M, Sharma C, Pandey J and Gupta P.K. found that the WTE generation potentials for RDF process range from 9 to 19, 8 to 18, 6 to 15 MW/day for Ghazipur (GL), Bhalswa (BL) and Okhla (OL) landfills respectively [10]

7. Literature Review

S.N	Paper Title	Journal	Author	Summary
1	Feasibility, health and economic impact of generation biogas from human excreta for the state of Tamil Nadu, India	Renewable and Sustainable Energy Reviews, 2017	Arunaachalam Muralidharan	Authors used human excreta to produce biogas, hence tackling both fuel and sanitation simultaneously by giving a short description about Tamil Nadu: Population density; Health & sanitation; culture; common village layout. A constant supply of bio fuel everyday can be provided with 13.1 million ton LPG replaced per year. Thus 75% of current consumption f LPG in India can be replaced by generation of biogas using human excretion. Challenges and concerns: like social taboo; risk of pathogens. Hence concluding biogas generation from human excreta can full fill future demand in India but require involvement of Government.

2	Quantification and Characterization of size-segregated Bioaerosols at municipal solid waste dumping site in Delhi	International Conference on Solid Waste Management, 5IconSWM 2016	Smita Agarwal et. al.	Delhi has four landfill sites: Bhalswa (North Delhi), Ghazipur (East Delhi), Okhla (South Delhi) and Narela-Bawana. And 3 are overflowing. The size-segregated aerosol were carried out during 2013-14 at Okhla landfill and collected on quart filters using an Andersen impactor sampler having 9 stages. And a low-volume handy sampler was also used to measure total cultural microbial concentrations. It was found that major concentration peaks during winter due to high PM concentration. Low during summer due to high temperature and solar radiation. Concentration (cfu/m3) = Number of colonies/flow rate (m^3/min)*sample duration (min). Bacterial range : 8.3*10^5 to 1.8*10^7 (cfu/m^3) Fungal range : 1.2*10^3 to 2.5*10^5 (cfu/m^3) Size distribution analysis shows that bacteria were mostly abundant in fine particle size. The result confirmed that close to the waste dumping ground bacterial and fungal aerosol may deposit in the nasal and oral cavities.
3	Biogas system in rural China : Upgrading from decentralized to centralized?	Renewable and Sustainable Energy Reviews, 2017	Qiu Chen et. al.	Author tried to balance the development of the two systems to enhance the contrasting strengths of them and reducing their weaknesses and discussed the construction of house hold base biogas digester from 1973 to 2014 and its output. Due to urbanization, large scale migration towards major city had lead to shortage in rural areas and thus increasing the labor cost. Hence life span of the household based biogas digesters is becoming shorter and shorter. The installation of centralized biogas plant in rural areas is beneficial since storage and transportation cost is fulfilled by the centralized body. Since energy demand of rural households is increasing constantly, the concept of centralized biogas plant is more successfully than decentralized plant. With the help of govt. policies and individual effort, enhancement of both centralized and decentralized biogas plant.
4	Mechanical characterization of municipal solid waste from two waste dumps at Delhi, India	Waste Management, 2017	Smita Agarwal et. al.	Author calculated the physical and mechanical properties of the emplaced municipal solid waste (MSW) located at Delhi. Ranging age – 2 to 12 years. Mechanical compressibility and shear strength were collected using a 300*300 mm direct shear box. Compression ratio varied from .11 to .17 Near the surface MSW consists of 60-80 % soil like or soil sized fraction. Solid waste specimens prepared with fibrous materials exhibited a continued increase of shear stress without any peak stress or failure even up to 55 mm displacement.

5	Municipal Solid Waste Management in India-Current State and Future Challenges : A Review	International Journal of Engineering Science and Technology (IJEST), 2012	Rajendra Kumar Kaushal et al.	Author estimated quantity and characteristics of municipal solid waste and its forecasting for successful management plan. This study analyses the changing trend in MSW quantities over the last 4 decades. Describing the relation between the amount of municipal waste and the GDP of the country. Higher the GDP, higher the quantity of waste produced. Studies have indicated that for every Indian Rs. 1000 increase in income the solid waste generation increases by one kilogram per month [23]. In low-income group of cities municipal bodies dispose MSW in low lying areas in the outskirt of the city and fill these areas one after the other haphazardly due to limited knowledge and awareness regarding contamination, waste reduction techniques and other aspects of MSW management. It suggests that MSW components like paper, plastic, glass are having the increasing trend from 4.1%, 0.7% and 0.4% respectively in 1971 to8.18%, 9.22% and 1.01% respectively in 2005.
6	Selection of appropriate biogas upgrading technology-a review of biogas cleaning, upgrading and utilization	Renewable and Sustainable Energy Reviews, 2015	Qie Sun et al.	Author discussed the various biogas technology, cleaning, upgrading and utilization for proper site-specific and dependence on biogas utilization. Biogas cleaning technology. Water scrubbing: water used as a solvent. Methane solubility is in water is much lower than that of CO and H_2. Water scrubbing can achieve a CH_4 purity of 80-99% depending on non-condensable gases like N_2 and O_2.Cryogenic Separation: Due to different condensing temperature of CH_4 and CO_2. CO_2 can be separated from CH_4 through condensation and distillation. This process needs to compress raw gas to high pressure up to 200bar, a large amount of energy is required, accounting to 5-10% of bio methane produced.
7	A Review on Organic Waste to Energy Systems in India	Bioresource Technology, 2017	Hiya Dhar et al.	The author identified the potential from the organic waste to energy conservation in India and its associated challenges. The biomass energy from firewood, crop residues, animal dung still used in rural parts of developing countries like India to meet the energy requirement. In India, the organic waste fraction varies between 40 to 60 % of the total solid waste streams. The calorific value of urban solid waste is 7.3 MJ/Kg and the moisture content is around 47% (Annnepu 2012). According to the database of Swachh Bharath Mission (SBM) only 21.45% of the MSW is treated and the remaining is still going to the landfills (SBM 2016). 2nd largest producer of paddy. India is producing 98 million metric ton (MT) of paddy and around 130 million MT of rice straw out of which 50% is used as fodder and remaining is thrown away.

8	Municipal solid waste: Generation, composition and GHG emissions in Bangalore, India	Renewable and Sustainable Energy Reviews, 2017	T.V. Ramachandra et. al.	Municipal solid waste in developing countries mainly consists of degradable materials (>70%), which plays a significant role in GHG emission in urban localities. Study revealed that per capita waste generation is about 91.01 g/day with per capita organic waste generation of 74 g/person/day. The household per capita waste generation was positively related with income and education levels, while negative related with family size. More than 90% of the MSW generated in India is disposed on land in unscientific and unacceptable way. Out of which major portion (70-75%) of municipal solid waste is organic. Most of the households (64%) in the study area have the facility of door to door collection of solid waste and about 78.34% of city population do not segregate the waste at source (households) Study also revealed that 82% of household wastes were organic. According to this study, the total organic waste generated from houses was 231.01 tons/year and total emission is about 604.8 t/year.
9	Does the construction of biogas plants affect local property values?	Economic Letters, 2017	Marco Modica	Although biogas is considered a renewable source of energy, the social acceptability of biogas plants is controversial due to resistance from communities. Author investigated this claim by using evidence from the housing market of Piedmont, where 167 biogas plants were opened between 2006 to 2015. The social acceptance of biogas plant is controversial due to resistance from local communities because of smell, heavy traffic and noise. Also opening of biogas plants had no effect on house prices.
10	A review on current status of municipal solid waste management in India	Journal of Environmental Sciences, 2015	Neha Gupta et. al.	Mismanagement of municipal solid waste can cause adverse environmental impacts, public health risk and other socio-economic problem. Author presented an overview of current status of solid waste management in India which can help the competent authorities responsible for municipal solid waste management and researchers to prepare more efficient plans. Treatment and disposal of municipal solid waste management. The composition of municipal solid waste at generation sources and collection points was determined on a wet weight basis and it consists mainly of a large organic fraction (40%–60%), ash and fine earth (30%–40%), paper (3%–6%) and plastic, and glass and metals (each less than 1%). The C/N ratio ranges between 800 and 1000 kcal/kg. In urban areas, the major fraction of municipal solid waste is compostable materials (40%-60%) and inerts (30%-50%).

11	Sustainable Municipal Solid Waste Management in India: A Policy Agenda	Procedia Environmental Sciences, 2016	Shyamala Mani and Satpal Singh	The Author reviewed the comprehensive review of the solid waste management system and most importantly highlight some major points of the government's policies and programs required to overcome the challenges of municipal solid waste management in India. 1. Greater emphasis on recycling is required. ULBs especially smaller ones can easily form cooperatives etc. with waste collectors initiatives and collectives for recycling rather than tie-ups with large companies for setting up incineration plants. 2. Recycling technologies for making structural from plastics or converting waste plastics and non-recyclables to Light Diesel Oil as mentioned in the Planning Commission report 2014 should be delineated and their standards prescribed. This is feasible in smaller ULBs also. 3. Much greater emphasis on ULBs working with waste pickers collectives and NGOs for setting up door to door collection of segregated waste, Material Recovery Facilities (MRFs) and secondary segregation while providing personal protective equipment (PPE) to them is required.
12	An empirical model for prediction of household solid waste generation rate – A case study of Dhanbad, India	Waste Management, 2017	Atul Kumar and S.R.Samadder	Accurate prediction of the quantity of household solid waste generation is very much essential for effective management of municipal solid waste (MSW).Author proposed, two models that established the relationships between the household solid waste generation rate and the socioeconomic parameters. The results of the two models showed that the coefficient of determinations (R2) were 0.782 for biodegradable waste generation rate and 0.676 for non-biodegradable waste generation rate using the selected independent variables. Validation of the developed models with a new set of data indicated a good fit for actual prediction purpose with predicted R2 values of 0.76 and 0.64 for biodegradable and non-biodegradable MSW generation rate respectively.
13	Qualitative determination of energy potential and methane generation from municipal solid waste (MSW) in Dhanbad (India)	Energy, 2017	Drake Mboowa et.al.	Methane generation from waste landfills is one of the biggest contributors to global warming. The purpose of this study was twofold: (i) to investigate methane concentration from Municipal Solid Waste (MSW) at three landfills in Dhanbad city, India (ii) to evaluate the amount of energy that could be recovered based on the MSW characteristics if it were to be incinerated. Results from MSW characterization revealed that the main component of Dhanbad MSW is organic waste, which made up to 75% of the waste by weight. Methane concentration and moisture content from Railway station (site

				1) and Memco-more (site 2 and site 3) measured as 140.53, 18.18 and 20.28 ppm methane/g waste and 25.49, 3.40 and 2.96% dry weight respectively. The calorific value for the waste samples ranged between 10.7 to 13.0 MJ/kg.
14	Innovation in solid waste management through Clean Development Mechanism in India and other countries	Process Safety and Environmental Protection, 2016	Aditee Potdar et. al.	Author discussed the potential of Clean Development Mechanism (CDM) for Municipal Solid Waste (MSW) management in India. About 350 MSW projects are implemented through CDM across 56 countries. The maximum MSW management CDM projects are implemented in China (102) followed by Brazil (45) and Mexico (28). About 22 CDM projects for MSW management are registered in India. The annual estimated emission reduction from these 22 projects is 1,467,371 ton CO2e/annum. The highest numbers of projects are registered in the state of Delhi. These CDM projects use technologies viz. refuse-derived fuel palletization, landfill capture, biogas generation, and composting. Though physical, operational regulatory, and socio-economic challenges exist for MSW management, India should make the best use of the opportunity that CDM offers and develop projects to benefit in terms of finance, technology and sustainable development.
15	Energy recovery from waste in India: An evidence-based analysis.	Sustainable Energy Technologies and Assessments, 2017	J.D.Nixon et. al.	The uptake of Waste-to-Energy (WTE) in India has not been successful and the majority of plants have failed to sustain operations. There is a lack of detailed on-the-ground research examining the causes of plant failures and the issues regarding the Waste to Energy supply chain. Local government officers, industry practitioners and academics involved in waste management in India were consulted. Quantitative data were collected on three case study plants: an incinerator, a gasification plant and a plant co-firing waste with coal. In comparison to the European incinerators, the WtE plants in India had a low capital cost (around 1–2 million €/MW), but total particulate matter emissions were a hundred times higher, ranging from 65-75 mg/Nm^3.
16	Municipal solid waste as a valuable renewable energy resource: a worldwide opportunity of energy recovery by using Waste-To-Energy Technologies.	9th International Conference on Sustainability in Energy and Buildings, SEB-17, 5-7 July 2017, Chania, Crete, Greece	Diego Moya et. al.	The generation rate of Municipal Solid Waste is expected to increase to 2.2 billion tonnes per year by 2025 worldwide. Author assessed the different waste-to-energy technologies developed to date. This work is divided into four groups: biological treatment of waste; thermal treatment of waste; landfill gas utilization; and bio refineries. Furthermore, integrated solid waste management systems with waste-to-energy technologies are studied and some worldwide examples are provided.

The improper disposal of municipal waste has a serious and dangerous impact on a wide range of areas. Garbage thrown in the street or in open spaces creates a public health hazard, while waste dumped near rivers, lakes and streams contaminates the water supply. Rubbish that is burned in the open rather than disposed of properly creates pollution and releases toxic fumes into the environment. Non-biodegradable materials thrown into open drains make their way into the sewerage system, clogging pipelines and damaging infrastructure. The hazards posed by the dumping of untreated hospital and industrial waste are even greater, with the release of pathogens and toxic compounds posing a grave threat not just to human life but also to plants and animals. Garbage dumped in the countryside is not simply an eyesore; entire landscapes are ruined and unique habitats for flora and fauna are lost.

There are various methods for mitigating the effect of municipal solid waste has been discussed in the above papers, the methods which are discussed in the papers above are landfills, incineration, composting, RDF, recycling and waste to energy plants. Some of these methods are being applied across various cities and better government policies across the country and better implication of these policies.

In order to evaluate existing knowledge gaps and barriers associated with institutional, social, legal and financial aspects, four major issues should be covered which are as follows: (1) lack of integrated sustainable waste management concepts in academic programmes, (2) lack of demonstration activities concerning SWM technologies and services, (3) lack of institutional capacity (referring to trained personnel, insufficient funds, policy constraints) and (4) insufficient access to know-how.

So the main obstacles in the success of these waste management methods are the government policies and their implications across the country.

8. Results and Discussions

While SWM was completely neglected in past and is now receiving some attention at the highest levels in several cities and states, many are lagging behind and several have not bothered to make any improvement at all. The national and state solid waste management missions need to be created to ensure that municipal authorities perform their obligatory duties regularly in compliance with MSW Rules 2000 within a predetermined time frame.

Though levels of SWM services in the country have started improving on account of active monitoring by the Supreme Court of India, the central and state pollution control boards and finance and technical support from proactive state governments there still is a long way to go. Save the formalization of the MSW Rules 2000, state action in this regard at many levels has been fairly uninspiring thus far. While MSW Rules 2000 is a watershed document in India's history of effective SWM, implementation issues still overwhelm the system.

A comprehensive nationwide program needs to be actively implemented keeping in mind possible future scenarios. Key individuals within the governing system and the bureaucracy need to be educated to the magnitude of the crisis and motivated to use their power to influence the system and appropriately channelize resources to actively promote effective and progressive SWM projects and practices.

From the above data we can see that waste management faces many problems in India to improve the situation the most important is to educate the population of India about potential negative effects of solid waste. The people should be educated to realize the importance of source segregation at generation point as biodegradables, inert and recyclable material for proper waste management. Moreover, manufacturing of non-recyclable polyethylene bags should be banned or research should be initiated to develop biodegradable polyethylene. Protection of groundwater contamination from leachate percolation from open dump/landfill site should be made compulsory. Appropriate technological solution should be adopted to achieve this goal.

9. Conclusion

The following conclusions were made**.**

1. There are various waste management techniques in India some of the main techniques used are landfills, incineration, composting, RDF, recycling and waste to energy plants.
2. Though there are these methods present but their implementation has not been done properly in our country.
3. The main concerns for the development of these methods across the country are the government policies and their implementation across the country.

References

[1] Hoornweg D. and Bhada-Tata P. What a waste: a global review of solid waste management, 2012.

[2] Ali M. (2006). Editorial: urban waste management as if people matter. Habitat International 30, p. 729–730.

[3] Zaman A.U., & Lehmann S. Urban growth and waste management optimization towards 'zero waste city'. City, Culture and Society 2, (2011). p. 177–187.

[4] Chan J.K.H. The ethics of working with wicked urban waste problems: The case of Singapore's Semakau Landfill. Landscape and Urban Planning 154, 2016, p. 123–131.

[5] Cheng H & Hu Y. Municipal solid waste(MSW) as a renewable source of energy: current and future practices in China. Bioresour Technol, 2010, p. 16–24.

[6] 6.Lino F.A.M., Ismail K.A.R. Energy and environmental potential of solid waste in Brazil. Energy Policy 39, 2011, p. 3496-3502.

[7] Singh R.P., Tyagi V.V., Allen T, Ibrahim M.H., Kothari R. An overview for exploring the possibilities of energy generation from municipal solid waste (MSW) in Indian scenario. Renew Sustain Energy Rev 15, 2011, p. 4797-4808.

[8] Census of India, 2011, General Population Table.

[9] Singh S & Mani S. Sustainable Municipal Solid Waste Management in India: A Policy Agenda, Procedia Environmental Sciences 35, 2016, p. 150-157.

[10] Chakraborty M, Sharma C, Pandey J, Gupta P.K. Assessment of energy generation potentials of MSW in Delhi under different technological options, Energy Conversion and Management 75, 2013, p. 249–255.

[11] https://www.sswm.info/sites/default/files/toolbox/UNEP%202002%20Sanitary%20Landfill.jpg, accessed on 08.11.2017, 20:00.

[12] Cleveland C.J., Morris C.G., Dictionary of energy. Elsevier, 2006.

[13] Tan S.T., Ho W.S., Hashim H., Lim J.S., Ho W.S., Lee C.T., Yan J.Y. Energy and emissions benefits of renewable energy derived from municipal solid waste: analysis of low carbon scenario in Malaysia. Applied Energy 136, 2016, p. 797-804.

[14] Manual on municipal solid waste management, Central Public Health and Environmental Engineering Organisation (CPHEEO), Ministry of Urban Development, New Delhi; 2000.

[15] MGBEPS (Malaysian Green and Blue Environment Protection Society), MSW Manual; 2012. [Cited 01.03.13]. <http://greenbluegroup.blogspot.in/2012/04/ incinerator-for-municipal-solid-waste.html>.

[16] Karagiannidis A, Bilitewski B, Tchobanoglous G, Themelis NJ, Wittmaier M, Tsatsarelis T. Waste to energy: on thermal treatment & energetic utilization of solid wastes. In: Golish TV, editor. Waste management research trends. New York: Nova Science Publisher, 2008, p. 105–163.

[17] http://www.epd.gov.hk/epd/sites/default/files/epd/english/environmentinhk/waste/prob_solutions/images/IWMF_shcematic_eng.jpg, accessed on 08.11.2017, 20:03.

[18] Roman P., Martinez M.M., Pantoja A. Farmer's compost handbook. Experiences in Latin America. Santiago FAO, 2015.

[19] Sikka P. Energy from MSW RDF Pelletization a Pilot Indian Plant; 2000. [Cited 01.03.13]. <http://www.environmental-expert.com/articles/energy-from-msw-rdf-pelletization-a-pilot-indian-plant-2080/view-comments>.

[20] http://www.climatetechwiki.org/sites/climatetechwiki.org/files/images/extra/composting_process.jpg, accessed on 20:04, 04/11/2017.

[21] http://biophysics.sbg.ac.at/waste/organic/step-1.jpg, accessed on 20:40, 04/11/2017.

[22] http://global.kawasaki.com/en/upload_images/img_rdf.jpg, accessed on 22:50,04/11/2017.

[23] Nixon J.D., Dey P.K., Ghosh S.K., Energy recovery from waste in India: An evidence-based analysis, In Sustainable Energy Technologies and Assessments, Volume 21, 2017, Pages 23-32, ISSN 2213-1388.

Talanta 187 (2018) 379–389

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Electrochemical genosensor based on template assisted synthesized polyaniline nanotubes for chronic myelogenous leukemia detection

Amrita Soni[a,b], Chandra Mouli Pandey[c], Shipra Solanki[a,b], R.K. Kotnala[b], Gajjala Sumana[b,*]

[a] Academy of Scientific and Innovative Research (AcSIR), CSIR-National Physical Laboratory, New Delhi 110012, India
[b] CSIR-National Physical Laboratory, New Delhi 110012, India
[c] Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

Keywords:
Cancer
Electrophoretic
Polyaniline nanotubes
Biosensor

ABSTRACT

This work reports a facile approach to synthesize polyaniline nanotubes (PANI-NT) by using manganese oxide as sacrificial templates. This template assisted polyaniline nanotubes (t-PANI-NT) were utilized as electrode material after deposition onto the indium tin oxide (ITO) coated glass substrates by using the electrophoretic technique. The structural, morphological and electrochemical characterizations of the t-PANI-NT show relatively better results compared to chemically synthesized PANI-NT (c-PANI-NT). Moreover, the t-PANI-NT/ITO electrode exhibits improved electron transfer coefficient ($\alpha = 0.63$) and charge transfer rate constant ($k_s = 0.05912\ s^{-1}$) in comparison to c-PANI-NT/ITO electrode ($\alpha = 0.56$ and $k_s = 0.06548\ s^{-1}$). The obtained t-PANI-NT/ITO electrodes have been further immobilized with biotinylated DNA sequence, specific to chronic myelogenous leukemia (CML) by using avidin–biotin as a cross-linking agent. Electrochemical impedance spectroscopy studies revealed that the genosensor displays linearity in wide range of target DNA concentration (10^{-6} to 10^{-16} M) with an outstanding differentiation ability and low detection limit of 10^{-16} M. The experimental results of this highly sensitive and specific genosensor with clinical samples of CML positive patients and control negative patients indicate its potential for clinical diagnostics.

1. Introduction

Nanostructured materials show several advantages, such as large surface area, high aqueous stability and the diverse conductive mechanism that lowers the interfacial impedance between the electrode and the electrolyte compared to their bulk counterpart [1,2]. Moreover to improve the performance of the nanostructured materials the control over the size and morphology at the nanoscale is very essential [3,4]. Synthesis using a pre-existing nanostructured template is one of the most effective strategies towards achieving high degree of synthetic control with respect to size, shape, composition and spatial arrangement of nanostructured materials, which usually benefits from the directing effect of the templates [5–8]. In order to synthesize ordered conducting polymers (CP) nanostructures, it is necessary to use close-packed assemblies of templates [9,10] as polymers are highly unstable at the nanometer scale. A variety of CP nanostructures have been developed, and their reversible interchange between redox states, excellent electrical conductivity, and enhanced stability make them suitable for applications in energy conversion, storage devices, sensors, actuators, and bio-devices [11,12]. Among various CP, polyaniline (PANI) has been considered as one the most studied conductive conjugated polymer due to its unique physical and chemical properties viz. high conductivity, good environmental stability, redox behavior and biocompatibility leading to many applications in gas sensing as well as in electrochromic and electrocatalytic devices [13–15]. PANI nanostructures, e.g., nanorods, nanofibers, nanospheres offer enhanced electrochemical performance by providing high interfacial area and have great potential for the development of miniaturized devices [16,17]. Moreover, it has been indicated that the tube-like structure of PANI could increase the electrode surface area and provide shorter and direct pathways for more charge transportation and ion diffusion which facilitate the electrode volume change in charge/discharge processes [18,19]. Therefore, a robust route to synthesize PANI nanotubes (PANI-NT) with well-controlled morphology and improved electrochemical performance can play a significant role for biomedical applications [20].

Chronic myeloid leukemia (CML) is a hematopoietic stem cell malignancy arises from reciprocal translocation of t(9;22) (q34;q11.2), results in a genetic recombination between the cellular ABL proto-oncogene on chromosome 9 and the breakpoint cluster region (BCR) gene

* Corresponding author.
E-mail address: sumanagajjala@gmail.com (G. Sumana).

https://doi.org/10.1016/j.talanta.2018.05.038
Received 24 February 2018; Received in revised form 7 May 2018; Accepted 8 May 2018
0039-9140/ © 2018 Elsevier B.V. All rights reserved.

on chromosome 22 [21,22]. The BCR/ABL chimerical fusion gene on 22q11 encodes for unrestrained tyrosine kinase which causes disorderly proliferation of leukocytes leading to blocking of apoptosis, genome instability and enlargement of the spleen, liver, and other organs, which plays a vital role in the pathogenesis of CML [23,24]. Currently, several useful and efficient clinical diagnostic techniques have been applied to recognize overexpressed, rearranged BCR/ABL fusion gene included fluorescence in situ hybridization, flow cytometry and real-time quantitative reverse transcription PCR, cytogenetics and southern blot [25,26]. But due to limitations of these methods, (time consumption, poor precision, expensiveness and error-prone nature), it becomes very significant to develop a new method that requires only a few handling steps, making the point-of-care diagnostics and monitoring of CML possible at an early stage [27–29]. In this context, electrochemical DNA biosensors have been recognized as a promising solution for simple, portable, rapid, and inexpensive sequence-specific detection of genetic diseases and other biological analysis [30–32]. Various electrochemical genosensors based on PANI nanostructures have been explored that provides a distinctive molecular platform for the immobilization of DNA in which PANI nanostructures act as molecular wires by connecting DNA molecules that provide controlled bio-molecular events [33–35].

In the present work, PANI-NT has been synthesized using both chemical and template route. The synthesized PANI-NT has been electrophoretically deposited onto indium tin oxide coated glass substrates. These PANI-NT films have been utilized for the fabrication of genosensor by immobilizing probe DNA (pDNA) sequence specific to CML, identified from the BCR–ABL gene. Further, these bioelectrodes were used for detection of CML target sequence both in synthetic DNA samples and clinical samples. To the best of our knowledge, the present work is the first report on the application of template-assisted PANI-NT based electrodes as the transducer platform for detection of cancer.

2. Experimental section

2.1. Reagents and chemicals

Potassium permanganate ($KMnO_4$), potassium dichromate ($K_2Cr_2O_7$), hydrochloric acid (HCl), avidin, 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC), N-Hydroxy succinimide (NHS), ethylene diamine tetra acetic acid (EDTA), potassium monohydrogen phosphate, potassium dihydrogen phosphate, camphor sulfonic acid (CSA), ammonium persulfate (APS), Aniline ($C_6H_5NH_2$) was distilled prior to polymerization. Indium–tin oxide (ITO) coated glass sheets with resistivity 30 Ω have been utilized. All other chemicals were of analytical grade and used without purification. For the preparation of aqueous solutions, deionized water (resistance 18.2 MΩ cm) was used. CML-specific oligonucleotide probe sequences (22 bases; UID no. 12859844) correspond specifically to P210 kDa BCR–ABL gene of Philadelphia chromosome [31], complementary, noncomplementary, one base-mismatch, two base-mismatch target sequences and all other reagents and solvents have been procured from Sigma-Aldrich Milwaukee, WI. DNA sequences used for the studies are as follows:

Probe DNA: biotin-5′-TGTCCACAGCATTCCGCTGACC-3′
Complementary DNA: 5′-GGTCAGCGGAATGCTGTGGACA-3′
One-base mismatch DNA: 5′-GCTCAGCGGAATGCTGTGGACA- 3′
Two-base mismatch DNA: 5′-GCTCAGCGCAATGCTGTGGACA-3′
Non-complementary DNA: 5′-TACTCGCAATAACGTGATCTCC-3′.

2.2. Processing of clinical samples

RNA from the blood samples of CML positive and negative patients was extracted by using QIAamp RNA blood mini-kit (Qiagen). Leukocytes were removed from lysed blood by selective erythrocyte lysis followed by repetitive homogenization and washing. Afterwards, high-capacity cDNA reverse transcription kit (Applied Biosystems) have been utilized for the reverse-transcription of separated RNA to complementary DNA (cDNA). The extracted cDNA was diluted 10 times in TE buffer and boiled in a water bath for about 5 min and then chilled in an ice bath followed by centrifugation at 12,000 rpm for about 10 min to get denatured single-stranded DNA. The aqueous layer of sample containing DNA was subjected to sonication at 24 kHz frequency for 15 min to break the long DNA strands into smaller fragments.

2.3. Apparatus and measurements

The structural and morphological characterizations of manganese oxide nanotubes (MnO_2 NT), PANI-MnO_2 nanotubes and PANI-NT have been investigated using transmission electron microscope (TEM, Hitachi Model, H-800) and scanning electron microscope (SEM, JEOLJSM-6700F, 10 kV). For elemental analysis of the materials synthesized, SEM equipped with an Energy Dispersive X-ray spectroscopy (INCA energy 250 EDX) analyzer has been used. X-ray diffraction (XRD, Rigaku Miniflex X-ray diffractometer) patterns were recorded in the 2θ range with a scan speed of 2°/min using Cu Kα radiation (λ = 0.1542 nm). Electronic conductivity was measured on compacted pellets using two probe conductivity method. Fourier Transform Infrared (FT-IR) spectroscopic investigations were performed using Perkin-Elmer spectrometer (model Spectrum BX) at 25 °C. Dynamic light scattering (DLS) measurements were performed at 25 °C with a Zetasizer Nano-ZS90 (Malvern Instruments) at a scattering angle of 90° to determine the zeta potential value of synthesized nanotubes. Contact angle (CA) investigations have been recorded using contact angle meter (Data Physics OCA15EC). Electrochemical measurements have been carried out on an Autolab potentiostat/galvanostat (Eco Chemie, Utrecht, The Netherlands) using a standard three-electrode cell system with PANI-MnO_2/ITO or PANI-NT/ITO as working electrode, Pt as an auxiliary electrode and Ag/AgCl as a reference electrode in Phosphate-buffered saline (PBS, 100 mM, pH 7.0, 0.9% NaCl) containing 5 mM of $[Fe(CN)_6]^{3-/4-}$.

2.4. Synthesis of MnO_2 nanotubes

MnO_2 NT were synthesized using earlier stated method with slight modifications [4]. In brief, 2.5 ml of concentrated HCl (37%) was added to 0.66 g of $KMnO_4$ dissolved in 75 ml of deionized water, and the solution was stirred for 15 min at room temperature. The solution was then transferred to a Teflon-lined stainless steel autoclave and was kept at a temperature of 140 °C in an electrical oven for 18 h [5]. After that, the obtained product was washed with DI water and ethanol several times till the pH value reaches to 7.0. The material was then dried in vacuum oven at 80 °C for 30 h.

2.5. Synthesis of PANI-NT by using templates

PANI-NT have been synthesized using MnO_2 as the sacrificial templates (t-PANI-NT) (Fig. 1(A)). For the synthesis, 500 mg of MnO_2 NT were dissolved in 50 ml of DI water with stirring for 15 min and further the solution was sonicated for 45 min [4]. Subsequently, 1 M HCl, 400 µl of aniline and 0.45 gm of $K_2Cr_2O_7$ was added to the solution with continuous stirring for 8 h [11], and then the mixture was allowed to stand overnight. The obtained product was collected and washed with ethanol and distilled water several times to remove the template (MnO_2 NT). The synthesized material was vacuum dried at 50 °C for 12 h. To delineate the reaction mechanism and to study the characteristics of PANI-MnO_2 NT composite for sensing applications, the material has been synthesized using the similar method at reduced reaction time interval. The products were washed carefully at mild conditions and dried at room temperature for 24 h.

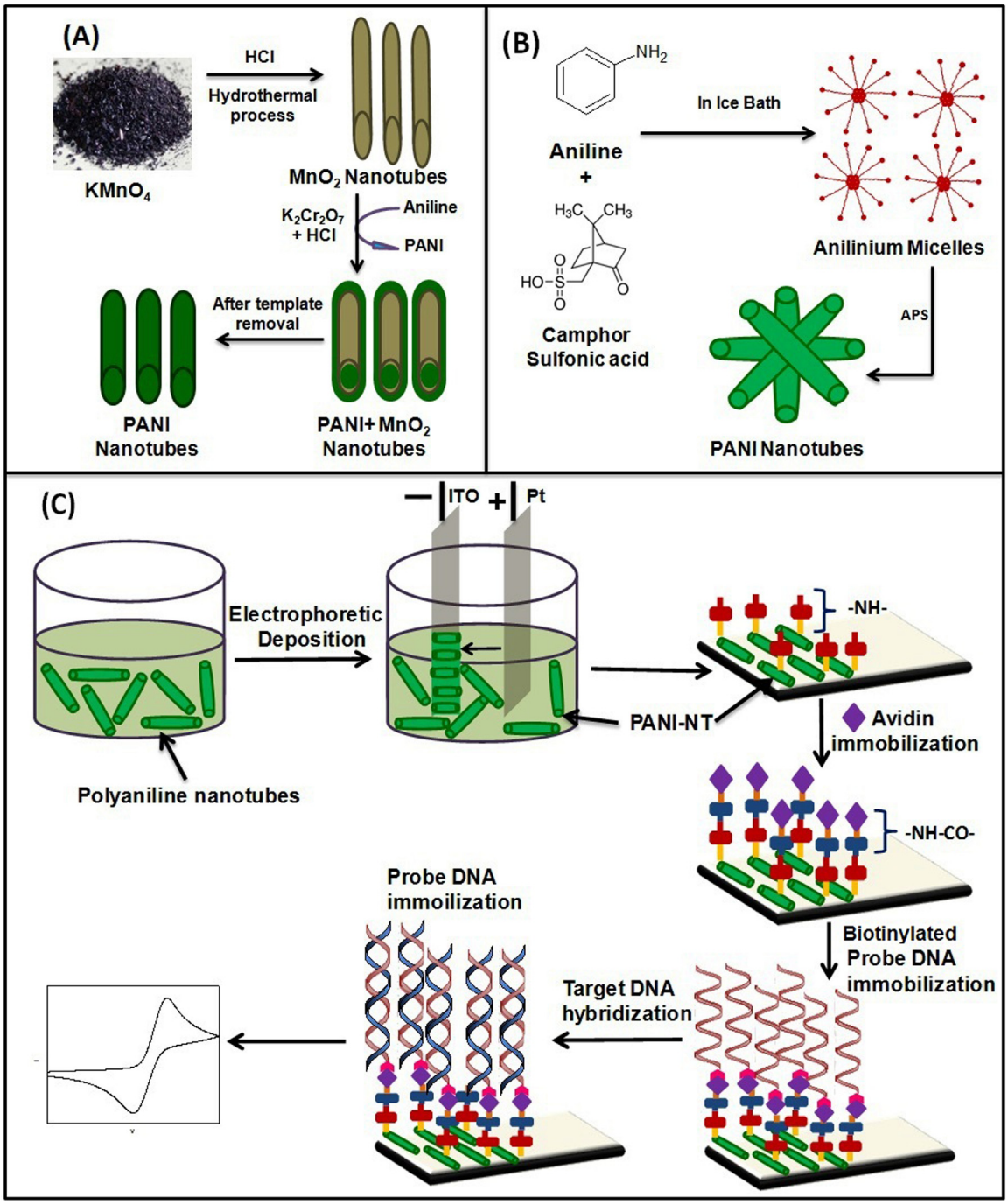


Fig. 1. Schematic illustration for the preparation of (A) PANI-NT by using MnO_2 nanotubes as reactive oxidative templates. (B) PANI-NT without using templates (C) Electrophoretic deposition of PANI-NT on ITO electrode and fabrication of pDNA/PANI-NT/ITO bioelectrode.

2.6. Synthesis of PANI-NT without using templates

Fig. 1(B) illustrates the preparation of the PANI-NT without using a template (c-PANI-NT) by chemical route [36]. In brief, 0.1 M of aniline was added to a pre-ice cooled solution of CSA (0.4 M) in 20 ml of DI water till the solution become transparent. The solution was cooled down in an ice bath, and 2 M cold aqueous solution of APS was added to the solution for oxidative polymerization [37]. The mixture was allowed to stand overnight in an ice bath. The precipitates were filtered and washed with distilled water and methanol several times and were dried in vacuum oven for 24 h at room temperature.

2.7. Electrophoretic deposition of PANI-NT and PANI-MnO_2 NT

Electrophoretic deposition (EPD) was carried out using a two electrode system, where ITO electrode was used as a cathode and thin platinum foil as an anode (Fig. 1(C)). For the deposition of t-PANI-NT onto ITO electrodes, 10 mg of t-PANI-NT was dispersed in 10 ml of ethanol and 20 V of voltage for 45 s was applied. Similarly, the deposition of c-PANI-NT on to ITO electrode was done at 100 V for 60 s by taking c-PANI-NT in a solution of 9.9 ml of acetonitrile and 100 µl of formic acid [36]. 10 mg of PANI-MnO_2 NT was dissolved in a solvent mixture of ethanol and water (1:1) and films of PANI-MnO_2 were deposited on ITO electrodes under the optimum conditions of 30 V for 60 s. The conditions were different for the two materials, due to the

variation in oxidant and dopant ($K_2Cr_2O_7$ and HCl for t-PANI-NT and APS and CSA for c-PANI-NT respectively) used for synthesizing the nanotubes, which further affect the solubility as well as depositing conditions of the material synthesized.

2.8. Fabrication of bioelectrodes for CML detection

For the fabrication of bioelectrodes, pDNA (22 base pairs) has been immobilized on the t-PANI-NT/ITO electrodes using avidin-biotin linkage. For this purpose, 20 µl of avidin was immobilized (2 h) on the t-PANI-NT/ITO platform by using EDC (0.2 M) and NHS (0.4 M) at 25 °C [38]. Subsequently, avidin-t-PANI-NT/ITO electrode was washed and incubated with biotinylated pDNA (20 µl, 1.0 mM) for 6 h in a humid chamber at 25 °C. These bioelectrodes were washed with Tris-HCl (10 mM) and EDTA (1 mM) buffer solution (TE, pH 8.0) several times to remove any unbound pDNA from the electrode surface. The fabricated pDNA/t-PANI-NT/ITO bioelectrodes were further utilized for the hybridization studies by incubating it with the target DNA sequences for 20 min in a humid chamber at 25 °C (Fig. 1(C)).

3. Results and discussion

3.1. Microscopic analysis of nanotubes

The TEM studies have been conducted to confirm the size and structure of the synthesized nanotubes. Fig. 2(A) illustrates the TEM image of MnO_2 NT synthesized using hydrothermal process, having a diameter in the range of 40–50 nm (inset of Fig. 2(A), showing MnO_2 NT with a diameter of 42 nm). MnO_2 NT exhibit smooth surface and are uniformly dispersed without formation of domains due to aggregation, which is confirmed by high-resolution transmission electron (HRTEM) microscopy image (Fig. 2(B)) of MnO_2-NT. The TEM image of c-PANI-NT (Fig. 2(C)) shows tubular structures with a diameter ranging from 90 nm to 130 nm (Fig. 2(D)), and along with nanotubes, nanofibers can also be seen in the image (Fig. 2(C)). The TEM images of PANI-MnO_2 (Fig. 3(A) and (B)) were recorded to analyze the growth process of PANI-NT on the surface of MnO_2 NT, in which, MnO_2 acts as template and oxidant or reducing agent at the same time. In Fig. 3(A), multiple layers of PANI wrapped around the surface of MnO_2 NT results in the formation of PANI-MnO_2 NT with increased diameter in the range 80–120 nm as compared to the MnO_2 NT. In Fig. 3(B), the TEM image of PANI-MnO_2 NT at higher magnification clearly reveals the coating of PANI on MnO_2 and presence of MnO_2 NT in the center can also be observed. Further, TEM studies reveal that with an increase in the reaction time the gradual growth of PANI-NT (Fig. 3(C) and (D)) and degradation of the MnO_2 template can be observed which leads to the formation of the tube-like structures of PANI (Fig. 3(C)). However, it can be seen that the t-PANI-NT exhibit the inner diameter in the range of 40–50 nm (Fig. 3(D)) which is comparable to the diameter of MnO_2 NT (Fig. 2(A)) with a wall thickness of about 14 nm. The dark areas (irregular projections) seen in Fig. 3(C) is may be due to multiple layers formed or folding of the edges of the t-PANI-NT as they are flexible and can be easily intertwined. It has been observed that t-PANI-NT are well dispersed and exhibited perfectly tubular structures (hollow) with open-ended tips (Fig. 3(C)) as compared with c-PANI-NT (Fig. 2(C)). This may be due to absence of directing effect provided by template, which also acts as driving force for the uniform synthesis of PANI-NT and without which the random accumulation of the PANI molecules will take place. Moreover, it is well known that the hydrogen bonding between the CP plays a very significant role in the development of uniform nanostructures, and disruption of which can cause formation of irregular shape of PANI-NT and further reduces the electrochemical performance of the material [16].

3.2. Characterization of PANI-NT and PANI-MnO_2 NT

Fig. 4(A) shows the XRD pattern of the MnO_2 NT. All the diffraction peaks indexed to tetragonal phase of α-MnO_2 (JCDPS No. 721982),

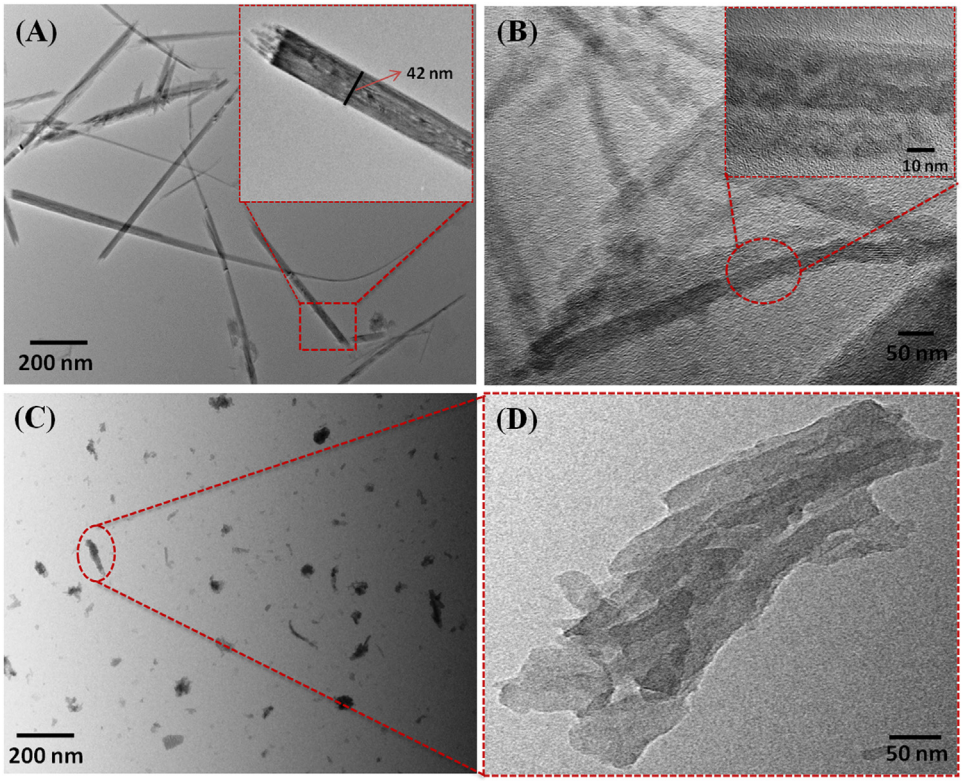


Fig. 2. (A) TEM image of MnO_2 NT (inset representing the diameter of nanotubes) (B) HR-TEM image of MnO_2 NT. TEM images of c-PANI-NT (C) at lower magnification (D) at higher magnification.

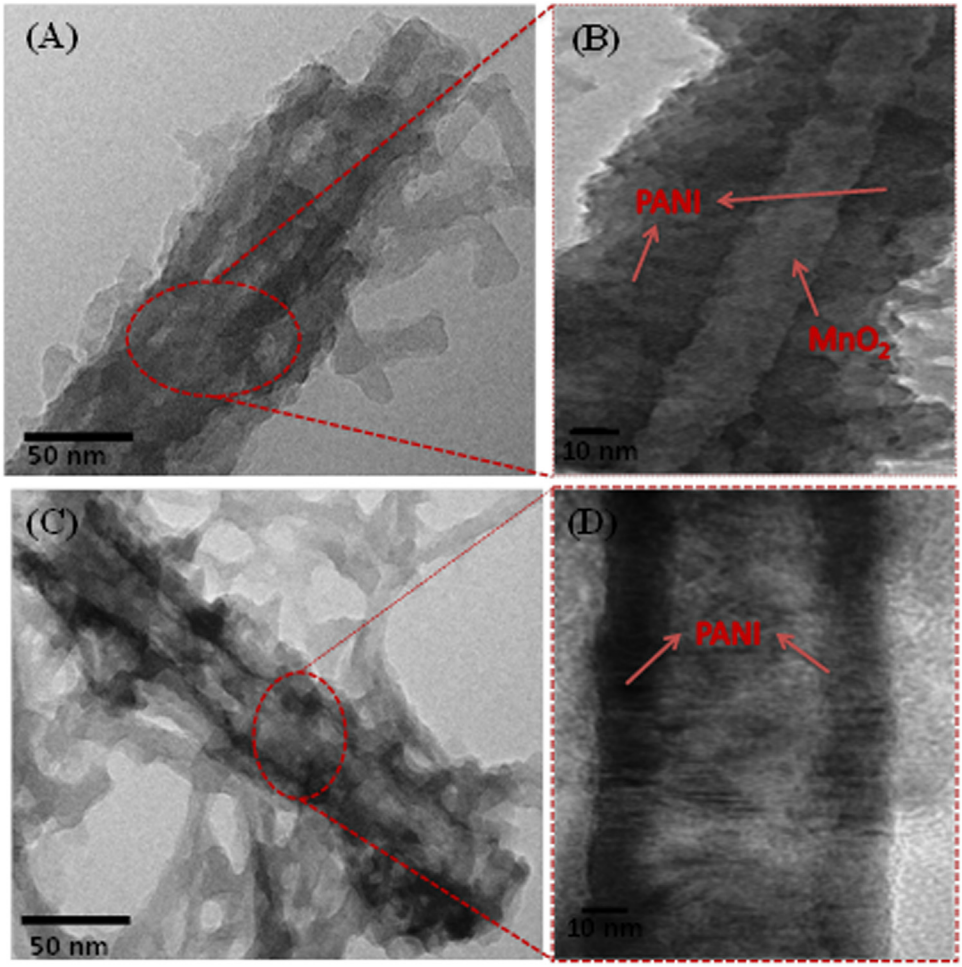


Fig. 3. TEM images of PANI-MnO_2 NT (A) at lower magnification (B) at higher magnification. TEM images of t-PANI-NT (C) at lower magnification (D) at higher magnification.

which reveals the crystalline behavior of the MnO_2 NT. After polymerization of aniline monomer, a layer of PANI has been formed on the MnO_2 NT (confirmed by the TEM image of PANI-MnO_2 (Fig. 3(A))), due to which the broadening of the peaks at $2\theta = 20.7^\circ$ and 26.3° can be observed, (Fig. 4(B)) representing the amorphous nature of the PANI-MnO_2 NT due to presence of PANI [39]. However, the peak at 20.7° disappears due to complete removal of the MnO_2 template, and only a single broad peak appears at $2\theta = 26.3^\circ$ (Fig. 4(C)) which shows controlled growth and the periodicity parallel to the polymer chains of PANI [5]. Fig. 4(D) shows the diffraction pattern of c-PANI-NT where a broad range of peaks from $2\theta = 18.3^\circ$ to 29.2° was observed, which signifies the ordered structure of the polymer chains having amorphous nature of the nanotubes synthesized. The XRD results in Fig. 4(C) and (D) inferred that the crystallinity of PANI nanostructures could be changed not only by the structure of the dopants used but also by controlling the morphology of PANI nanostructures [11].

The electrical conductivities of the synthesized PANI-MnO_2 NT, t-PANI-NT and c-PANI-NT, were carried out using two probe method at room temperature. The conductivity value of t-PANI-NT (4.16×10^{-2} S/cm) shows an improvement of about two orders of magnitude compared to the conductivity of PANI-MnO_2 NT (1.72×10^{-6} S/cm), which is probably due to the non-conducting nature of MnO_2, that provides more resistance to charge flow [5]. Moreover, the conductivity value of t-PANI-NT was also found higher as compared to c-PANI-NT (2.1×10^{-3} S/cm). This increased value indicates that the conductivity of the PANI nanostructures not only depends on the strength of the doping agent used and ionic bonding with the imine nitrogen but also on the synthesis route used and morphology of nanostructures, which further offer low potential energy to electron transfer and provide high electrochemical stability [40].

EDX studies were conducted to distinguish the chemical composition of material synthesized. EDX spectrum in the Fig. S1(A) displays the peaks corresponding to Mn and O, confirming the presence of MnO_2. However, the EDAX spectrum of PANI-MnO_2 (Fig. S1(B)) shows additional peaks of C and N, which correspond to the polymerization of the aniline on the MnO_2 templates (the Cl peak is present due to doping by HCl). Whereas, no Mn peaks were observed after the removal of Mn (Fig. S1(C)), indicating complete removal of the MnO_2 templates and the successful preparation of the t-PANI-NT. The inset of the each Fig. S1(A)–(C) describes the % weight of the each element present in the products.

The zeta potential values of the synthesized PANI-MnO_2, t-PANI-NT and c-PANI-NT have been measured via dynamic light scattering. Smoluchowski method was used to determine the total charge on the surface of prepared nanostructures (Fig. S2) [41]. The PANI-MnO_2 NT exhibited a zeta potential value of 13.4 mV (Fig. S2(A)) whereas, t-PANI-NT had a value of 31.6 mV (Fig. S2(B)) and c-PANI-NT had 28.5 mV (Fig. S2(C)) of charge on the surface. The high zeta potential of t-PANI-NT indicates stability of these nanostructures, which is the required characteristic for the electrophoretic deposition.

3.3. Morphological studies of prepared electrodes and bioelectrode

The surface morphological studies of the synthesized MnO_2 NT, t-

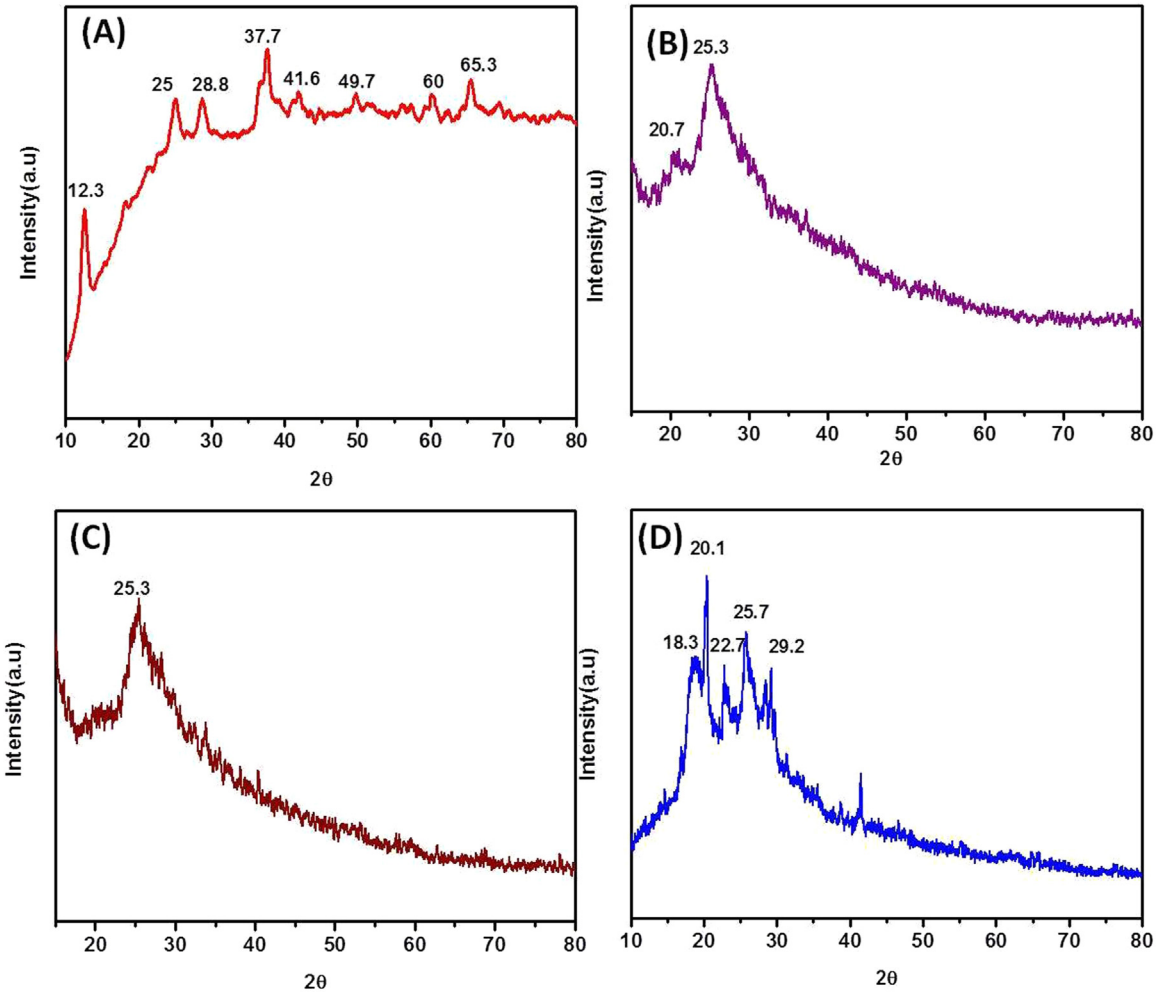


Fig. 4. X-ray diffraction patterns of (a) MnO_2 NT (b) PANI-MnO_2 NT (c) t-PANI-NT (d) c-PANI-NT.

PANI-NT/ITO electrode and c-PANI-NT/ITO electrode and pDNA/t-PANI-NT/ITO bioelectrode have been investigated using SEM. Fig. 5(A) demonstrate the SEM image of MnO_2 NT synthesized by hydrothermal reaction. The SEM image of c-PANI-NT/ITO electrode (Figs. 5(B) and S3(A)) indicates the agglomeration of the nanotubes that leads to the formation of clusters on the surface (inset Fig. 5(B)). Figs. 5(C) and S3(B) shows the morphology of t-PANI-NT/ITO electrode, where the nanotubes formed are in isolation without any agglomeration, maintaining their nanotubular structure and also exhibit the dense particle distribution. After incubation of the pDNA on t-PANI-NT/ITO electrode (Fig. 5(D)) the change in the morphology can be correlated to the immobilization of pDNA. The uniform morphology is due to even distribution of t-PANI-NT on the surface of ITO electrode that may provide a favorable environment for the immobilization of pDNA in an oriented manner [42,43].

3.4. FT-IR spectroscopic and contact angle measurements

Fig. S4 shows the FT-IR spectrum of the PANI-MnO_2/ITO, t-PANI-NT/ITO, c-PANI-NT/ITO electrodes and pDNA/PANI-NT/ITO bioelectrode. In the FT-IR spectrum of PANI-MnO_2/ITO electrode (Fig. S4(A)), the peaks at 524 and 722 cm^{-1} correspond to Mn–O stretching vibrations which are not present in the FT-IR spectrum of t-PANI-NT/ITO electrode (Fig. S4(B)) which clearly reveals the absence of MnO_2 in t-PANI-NT. Peak centered at 1364 cm^{-1} (Fig. S4(A)) is due to O–H stretching vibrations of Mn atom and vibrational peaks centered at 1581 and 1467 cm^{-1} are attributed to the stretching peaks of quinone and benzene ring deformations of PANI respectively. Whereas the FT-IR spectrum of t-PANI-NT (Fig. S4(B)) shows characteristic peaks of quinoid and benzenoid vibrations, that get shifted due to the removal of the template and are at 1582 and 1494 cm^{-1} respectively. Moreover, IR spectrum of t-PANI-NT shows (Fig. S4(B)) peak at 1295 cm^{-1} which corresponds to C–N stretching vibration of secondary aromatic amine, whether the peak at 1140 cm^{-1} corresponds to the characteristic N = Q = N stretching mode of the quinoid unit of PANI, and 801 cm^{-1} peak is assigned to C–H out-of-plane bending of 1,4 disubstituted benzene. However, in FT-IR spectrum of PANI-MnO_2/ITO electrode (Fig. S4(A)) very weak signals corresponding to the vibrational modes of PANI can be observed which are due to the presence of MnO_2 template [44]. The FT-IR spectrum of c-PANI-NT/ITO electrode (Fig. S4(C)), shows peaks at 1571 and 1495 cm^{-1} that correspond to stretching vibrations of the quinoid and benzenoid ring respectively. The band observed at 1290 cm^{-1} corresponds to C–N stretching vibration with aromatic conjugation and peak at 1160 cm^{-1} attributed to aromatic C–H in-plane bending while the peak at 830 cm^{-1} is due to bending vibrations of C-H bond in 1, 4 disubstituted benzene. The bands at 1034 and 595 cm^{-1} are assigned to absorption of -SO_3 H group indicating doping of PANI with CSA [13]. Fig. S4(D) shows the FT-IR spectrum of pDNA/t- PANI-NT/ITO bioelectrode, where the absorption bands at 1068 and 1231 cm^{-1} are due to the asymmetric stretching of P–O–C vibration and stretching vibration of P = O of the phosphoric acid group respectively. Whereas, the stretching vibration of C = C in the purine and pyrimidine attributed to carbonyl group was observed at 1550 and 1603 cm^{-1} respectively.

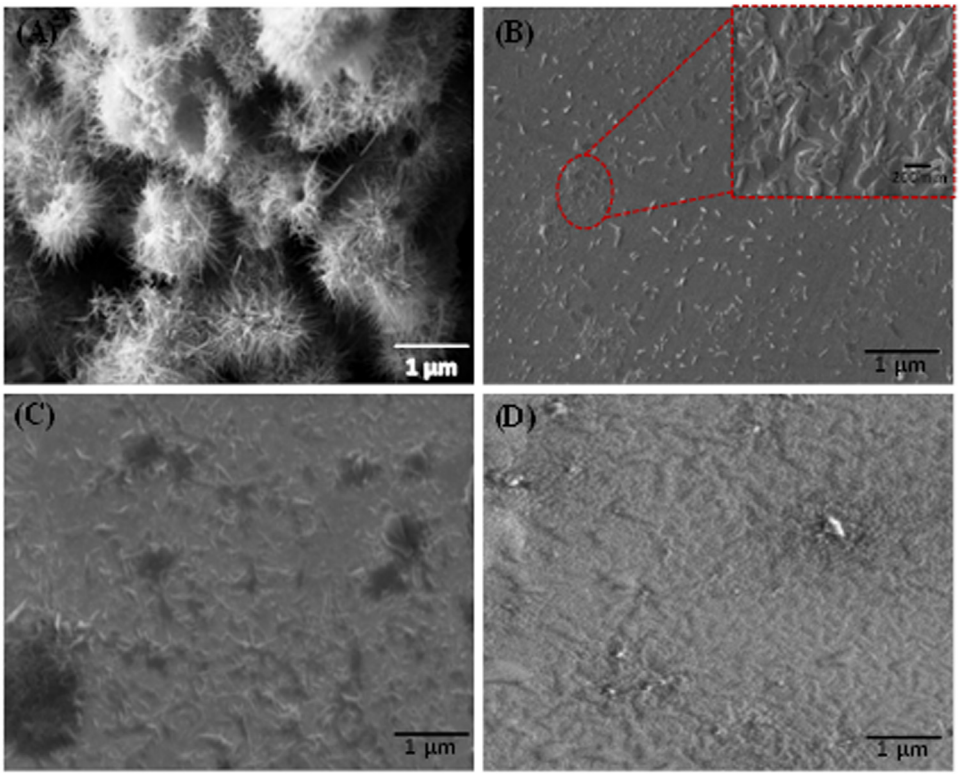


Fig. 5. SEM images of (A) MnO_2 NT (B) c-PANI-NT/ITO electrode (inset representing the image at higher magnification) (C) t-PANI-NT/ITO electrode (D) pDNA/t-PANI-NT/ITO bioelectrode.

Contact angle (CA) studies have been carried out using the sessile drop method to investigate the hydrophilicity/hydrophobicity of the electrodes and bioelectrodes. CA of pristine ITO electrode was measured to be 69.8° (Fig. S5(A)), which increases to 84.2° for PANI-MnO_2/ITO electrode (Fig. S5(B)). This increase is may be due to the introduction of hydrophobic groups on the surface of ITO electrode. The value CA for c-PANI-NT/ITO electrode was found to be 59.1° (Fig. S5(C)) and for t-PANI-NT/ITO electrode was 30.3° (Fig. S5(D)). The subsequent decrease in the value of CA can be directly related to more availability of hydrophilic functional groups (–NH_2) on the surface of the electrodes. The decrease in CA to 13.2° after the immobilization of pDNA (Fig. S5(E)) on t-PANI-NT/ITO electrode surface is due to the presence of hydrophilic negatively charged phosphodiester backbone of pDNA, which in turn can be related to incubation of pDNA on the electrode surface [42].

3.5. Electrochemical characterizations

To examine the electro-active behavior of PANI-MnO_2/ITO electrode, t-PANI-NT/ITO electrode, c-PANI-NT/ITO electrode and pDNA/t-PANI-NT/ITO bioelectrode, electrochemical studies have been carried out in phosphate buffer saline (100 mM, pH 7.0, 0.9% NaCl) containing 5 mM of $[Fe(CN)_6]^{3-/4-}$. An increase in the cyclic voltammetry peak current was observed for t-PANI-NT/ITO electrode (1.01 mA, Fig. S6(A); curve (iv)) compared to PANI-MnO_2/ITO electrode (635.30 µA, Fig. S6(A); curve (ii)), c-PANI-NT/ITO electrode (863.23 µA, Fig. S6(A); curve (iii)) and ITO (451.28 µA, Fig. S6(A); curve (i)) electrode. The higher peak current for t-PANI-NT/ITO electrode compared to PANI-MnO_2/ITO electrode, and c-PANI-NT/ITO electrode is due to the better transportation and diffusion of electrolyte ions in the electrochemical cell. Moreover, the presence of less conducting MnO_2 core also limits the access of electrolyte to the entire composite surface. However, upon immobilization of pDNA on to t-PANI-NT/ITO electrode decrease in the value of peak current (716.26 µA, Fig. S6(B); curve (ii)) has been observed which further decreases after hybridization with the target DNA (604.00 µA, Fig. S6(B); curve (iii)). This decrease in the current is due to the presence of a long chain of DNA that slows the electron transfer across the bioelectrode.

To investigate the various kinetic parameters, the cyclic voltammograms of the PANI-MnO_2/ITO, c-PANI-NT/ITO, and t-PANI-NT/ITO electrodes (Fig. S7) were recorded as a function of scan rate (10–300 mV s^{-1}). It has been observed that by increasing scan rate, the oxidation peak gets shifted towards more positive potential and reduction peak gets shifted in the reverse direction, exhibiting a quasi-reversible process. The redox peak current values of these electrodes follow a linear relationship with respect to the square root of scan rate, indicating the existence of diffusion-controlled electrochemical reaction (inset Fig. S7(A)–(C)) [45]. Both the redox peak potential values were found to be linearly dependent on the log of the scan rate which agrees with the Laviron's theory [42].

$$E_{pa} = E^o + Xln[(1-\alpha)Fv/RTk_s] \quad (1)$$

$$E_{pc} = E^o + Yln[\alpha Fv/RTk_s] \quad (2)$$

$$lnk_s = \alpha ln(1-\alpha) + (1-\alpha)ln\alpha - ln(RT/nFv) - \alpha(1-\alpha)nF\Delta Ep/RT \quad (3)$$

The value of electron transfer coefficient (α), for n number of electrons, was measured from the slopes of straight lines plotted, 2.303RT/(1 − α)nf and −2.303RT/αnf for the anodic peak potential (E_{pa}) and cathodic peak potential (E_{pc}) versus *ln v* respectively. The values of α and K_s for different electrodes (PANI-MNO_2/ITO, t-PANI-NT/ITO and c-PANI-NT/ITO) are represented in the Table 1. By using Randles–Sevcik equation [42]

$$Ip = (2.99 \times 10^5)\alpha^{1/2}n^{3/2}AC^{1/2}v^{1/2} \quad (4)$$

The diffusion coefficient (D) was determined by a concentration gradient of $[Fe(CN)_6]^{3-/4-}$ ions between bulk and interface. Here, C is the molar concentration of $[Fe(CN)_6]^{3-/4-}$, n is the number of transferred

Table 1
Kinetic parameters calculated for the PANI-MnO_2/ITO, t-PANI-NT/ITO and c-PANI-NT/ITO electrodes.

Name of electrode	Electron transfer coefficient, α	Charge transfer rate constant,(k_s/s^{-1})	Effective surface area,(A_{eff}/mm^2)	Diffusion coefficient, D(cm^2 s^{-1})
PANI-MnO_2/ITO	0.29	0.08396	0.497	3.01×10^{-14}
c-PANI-NT/ITO	0.56	0.06548	0.599	3.26×10^{-14}
t-PANI-NT/ITO	0.63	0.05912	0.711	4.59×10^{-14}

electrons for the redox reaction, and v is the scan rate (50 mV s^{-1}). By using the calculated diffusion coefficient, the electroactive surface area (A) of the different electrodes has been calculated using Eq. (5).

$$A = S/(2.99 \times 10^5)\alpha^{1/2}n^{3/2}CD^{1/2} \quad (5)$$

The various kinetic parameters calculated for the different electrodes have been summarized in Table 1. The enhanced kinetic parameters of t-PANI-NT/ITO electrode in comparison to PANI-MNO_2/ITO electrode are due to the restricted diffusion and transportation of the electrolyte ions as MnO_2 is present in the center and prevents the effective use of the inner surface of the PANI-NT. Moreover, the kinetic parameters of t-PANI-NT/ITO electrodes (Table 1) are also better than c-PANI-NT/ITO electrodes, which is due to more ordered arrangements of t-PANI-NT onto ITO electrode providing more surface area and improved charge transfer.

Electrochemical impedance spectroscopy (EIS) is an effective technique for studying the electronic characteristics of the electrodes before and after surface modification at electrode-solution interface. The characteristic EIS plots of the different electrodes (PANI-MNO_2/ITO,t-PANI-NT/ITO, and c-PANI-NT/ITO) are illustrated in Fig. 6(A), and an equivalent circuit (Randles model) is mentioned in the inset of Fig. 6(A). The R_{ct} is the interfacial electron-transfer resistance, C_{dl} corresponds to the double-layer capacitance, R is the ohmic resistance, and Z_w is the Warburg impedance. EIS was carried out in the frequency range 0.01–10^5 Hz, and phosphate buffer (100 mM, pH 7.0, 0.9% NaCl), $[Fe(CN)_6]^{3-/4-}$ was used as a redox marker. In the Nyquist impedance plot (Z_{re} *versus* Z_{im}), semicircle portion observed at higher frequencies represents the electron-transfer-limited process, and the lower frequency portion represents the diffusion-limited transport of the redox ions in the electrolyte to the electrode interface. Change in the diameter of the semicircle has been utilized to see the corresponding change in the electron-transfer resistance (R_{ct}) of the electrodes [46]. The R_{ct} value of the t-PANI-NT/ITO electrode (554 Ω) (Fig. 6(A); curve (i)) was lower in comparison to the c-PANI-NT/ITO electrode (670 Ω) (Fig. 6(A); curve (ii)) and the PANI-MnO_2/ITO electrode (735 Ω) (Fig. 6(A); curve (iii)). This decrease in the R_{ct} value for t-PANI-NT/ITO electrode is due to the higher conductivity value of t-PANI-NT, which facilitates the fast charge transfer capability and perfectly tube-like structures, higher protonation level of t-PANI-NT provides a more exposed surface to carryout electrochemical reactions. Further, after immobilization of probe DNA sequence on the t-PANI-NT/ITO electrode there was an increase in the R_{ct} value to 1115 Ω (Fig. 6(B); curve (ii)), suggesting the presence of negatively charged (phosphate group) pDNA on the electrode surface which repels negatively charged redox marker $[Fe(CN)_6]^{3-/4-}$. After incubating the pDNA/t-PANI-NT/ITO electrode with complementary target DNA sequence (cDNA), there was an increase in the R_{ct} value (Fig. 6(B); 3590 Ω; curve (iii)) which is may be due to the introduction of more negatively charged moieties (hybridization) onto the bioelectrode surface.

3.6. Electrochemical response studies of the bioelectrode

For sensitivity studies of the pDNA/t-PANI-NT/ITO bioelectrode, we have used EIS as a suitable investigation technique (Fig. 7(A)). It was observed that the R_{ct} value increases with increasing the cDNA concentration, indicating more hybridization reaction of pDNA with target DNA. The linear increase in the R_{ct} value is due to the increased negative charge on the bioelectrode that inhibits the electron-transfer kinetics. The analytical signal (in terms of $\Delta R_{ct} = R_{ct}c_{DNA} - R_{ct}p_{DNA}$) exhibited a linear relationship with the logarithmic value of the complementary target DNA concentration ranging from 10^{-16} to 10^{-6} M for pDNA/t-PANI-NT/ITO bioelectrode (Fig. 7(B); curve (i)) and from 10^{-13} to 10^{-6} M for pDNA/c-PANI-NT/ITO bioelectrode (Fig. 7(B); curve (ii)). The pDNA/t-PANI-NT/ITO bioelectrode exhibited linear regression coefficient (RC) of 0.9980 with standard deviations of 0.62 and for pDNA/c-PANI-NT/ITO bioelectrode the value of RC and standard deviations were found to be 0.9968 and 0.76, respectively. The imprecision of the data indicated by error bars and was found to be < 4% for pDNA/t-PANI-NT/ITO bioelectrode and < 6% for pDNA/c-PANI-NT/ITO. The analytical signal (ΔR_{ct}) is represented by following equation [47]:

$$\Delta R_{ct} = 73.030 logC + 1028.957 \quad (6)$$

The detection limit was calculated using the equation (3σ/sensitivity)[47], and the value was found to be 1×10^{-16} M (Fig. 7(B); curve (i)) for pDNA/t-PANI-NT/ITO and 1×10^{-13} M (Fig. 7(B); curve

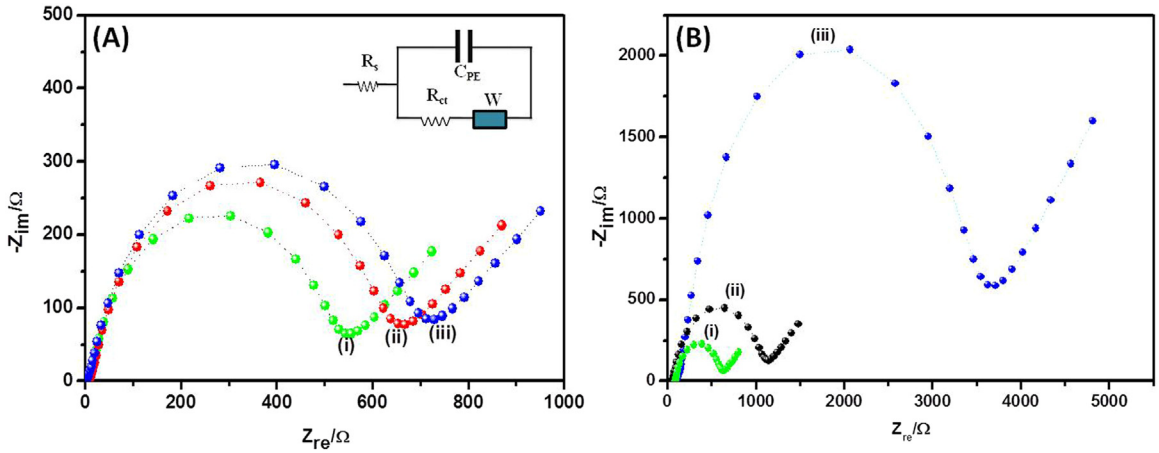


Fig. 6. Nyquist diagram for the Faradic impedance of (A) (i) t-PANI-NT/ITO electrode (ii) c-PANI-NT/ITO electrode (iii) PANI-MnO_2/ITO electrode (B) (i) t-PANI-NT/ITO electrode (ii) pDNA/t-PANI-NT/ITO bioelectrode (iii) pDNA/t-PANI-NT/ITO bioelectrode after hybridization with complementary target DNA sequence.

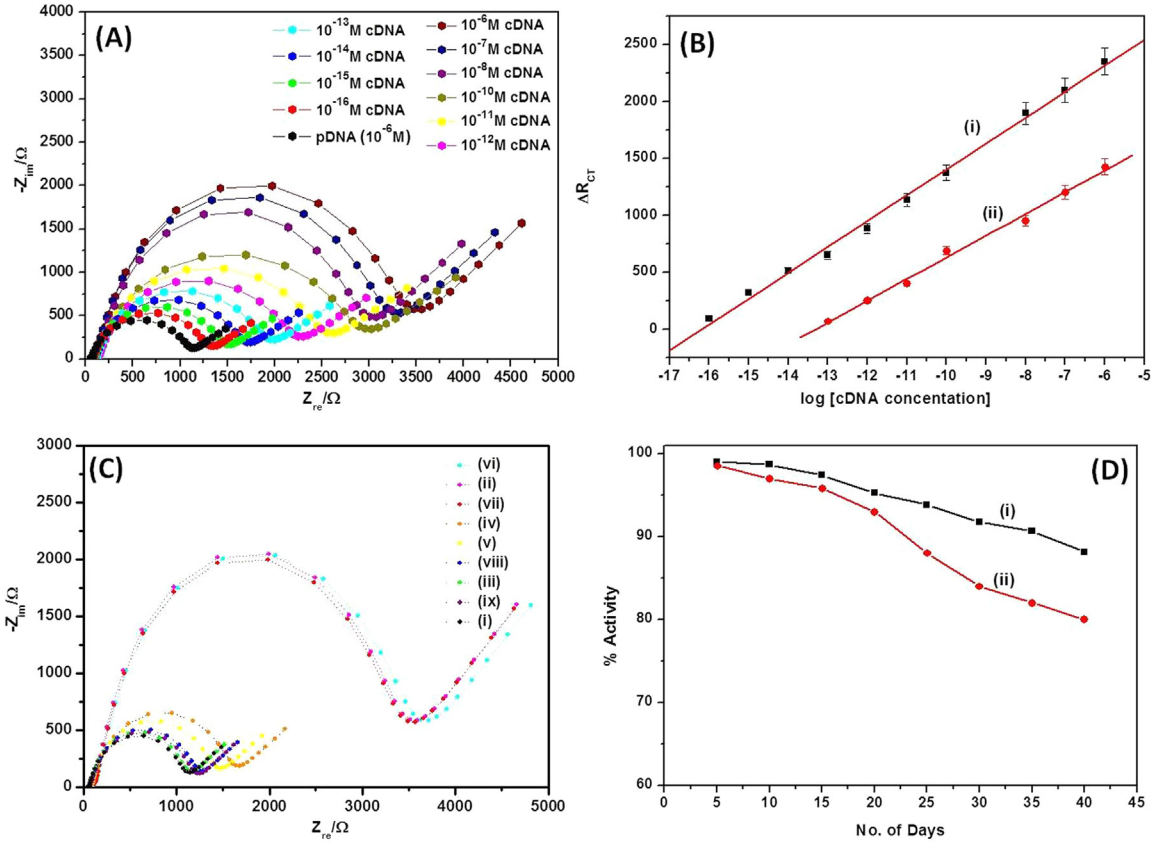


Fig. 7. (A) EIS response of pDNA/t-PANI-NT/ITO bioelectrode as a function of cDNA concentration (10^{-16} to 10^{-6} M) (B) Linearity relation curve between ΔR_{ct} and log of concentrations of cDNA for (i) pDNA/t-PANI-NT/ITO bioelectrode (ii) pDNA/c-PANI-NT/ITO bioelectrode (C) EIS response of pDNA/t-PANI-NT/ITO bioelectrode (i) with (ii) complementary (iii) non-complementary (iv) one-base mismatch (v) two base mismatch (vi) CML positive patient sample1 (vii) and sample 2 (viii) CML negative patient sample 1 (ix) and sample 2 (D) Stability of the (i) t-PANI-NT/ITO electrode and (ii) c-PANI-NT/ITO electrode after 40 days.

(ii)) for pDNA/c-PANI-NT/ITO bioelectrode, where σ is the standard deviation of the electrode without hybridization. The improved detection limit of the pDNA/t-PANI-NT/ITO bioelectrode in comparison to the pDNA/c-PANI-NT/ITO bioelectrode is due to the ordered arrangement and appropriate orientation of the pDNA onto t-PANI-NT/ITO electrode.

3.7. Selectivity and real sample analysis of bioelectrode

Selectivity studies of the prepared pDNA/t-PANI-NT/ITO bioelectrode were conducted by incubating it with different target DNA sequences (complementary, noncomplementary, one base mismatch and two base mismatch). When the pDNA/t-PANI-NT/ITO bioelectrode (Fig. 7(C); R_{ct} = 1115 Ω, curve (i)) was incubated with complementary target DNA sequence, there was a markable increase in the R_{ct} value (Fig. 7(C); 3590 Ω; curve (ii)) with respect to pDNA, which shows complete hybridization with the probe sequence. After incubation with non-complementary DNA sequence, no significant change in the R_{ct} value (Fig. 7(C); curve (iii)) was observed in comparison to the pDNA/t-PANI-NT/ITO bioelectrode indicating non-hybridization at the electrode surface due to mismatching of the DNA base pairs. However, after incubating the pDNA/t-PANI-NT/ITO bioelectrode with one base and two base mismatch DNA sequences, a slight increase in the R_{ct} value to 1664 Ω (Fig. 7(C); curve (iv)) and 1497 Ω (Fig. 7(C); curve (v)) for one base mismatch and two base mismatch respectively was obtained. This slight increase in the R_{ct} value indicates the partial hybridization of the target DNA with pDNA. This shows that the fabricated bioelectrode is even able to distinguish a single nucleotide variation during hybridization, in target DNA sequence.

3.8. Validation with clinical samples

Biosensor response studies with clinical patient samples have been recorded by measuring the change in the R_{ct} value after incubating the bioelectrode with different positive and negative patient samples of CML and are shown in Fig. 7(C). It was observed that the R_{ct} values (Fig. 7(C); curve (vi and vii)) of the CML positive patient samples are similar to the R_{ct} value of cDNA, indicating the complete hybridization however, no significant change in the R_{ct} value (Fig. 7(C); curve (viii and ix)) was seen when pDNA/t-PANI-NT/ITO bioelectrode was incubated with CML negative patient samples.

The change in the R_{ct} values for cDNA and CML positive and negative patient samples after incubating with pDNA/t-PANI-NT/ITO bioelectrode was also statistically analyzed by Analysis of Variance (ANOVA) by using Mann Whitney *t*-test for multiple comparisons (Fig. S9). Statistical significance threshold was set at 0.05 ($p < 0.05$). On hybridizing the pDNA/t-PANI-NT/ITO bioelectrode (Fig. S9(i)) with complementary target DNA the change in the R_{ct} value was found significant ($n = 3$, $p < 0.05$) (Fig. S9(ii)) in comparison to the pDNA/t-PANI-NT/ITO bioelectrode. Whereas, we did not observe a significant difference in the R_{ct} value for CML negative patient samples and in comparison to pDNA/t-PANI-NT/ITO bioelectrode ($n = 3$, $p > 0.05$) (Fig. S9(v) and (vi)). However, R_{ct} value for the CML positive patient samples were found to be statistically different, when compared with the R_{ct} value of bioelectrode ($n = 3$, $p < 0.05$) (Fig. S9(iii) and (iv)). The results indicate the validity of genosensor for detecting clinical patient samples.

Table S1 shows the comparative analysis of the results obtained from RT-PCR and biosensing assay for four different samples of both CML-positive and negative patients, having BCR-ABL transcript levels

Table 2
Comparison of the response characteristics of t-PANI-NT/ITO electrodes with other polyaniline based nanocomposites.

S. No.	Electrode	Method of immobilization	Electrochemical technique	Detection limit (M)	Detection range (M)	Hybridization time	Ref.
1.	Polyaniline nanowire deposited on GCE	EDC/NHS	SWV	1.0×10^{-15}	2×10^{-15}–8×10^{-13}	60 min	[49]
2.	BdNG-avidin-nsPANI/ITO	Biotin-avidin coupling	DPV, CV	0.5×10^{-15}	1×10^{-16}–1×10^{-6}	60 s	[38]
3.	PANI/rBi_2S_3 nanocomposite on carbon paste electrode	Covalent bonding	EIS, CV	4.37×10^{-16}	1×10^{-15}–1×10^{-11}	40 min	[50]
4.	PPy–PANi–Au nanocomposite	Covalent bonding	EIS, CV	1×10^{-13}	1×10^{-13}–1×10^{-6}	90 min	[51]
5.	Polyaniline/Pristine Graphene on GCE	Covalent bonding	EIS	0.01×10^{-12}	1×10^{-10}–1×10^{-6}	20 min	[52]
6.	t-PANI-NT/ITO	Biotin-avidin coupling	EIS, CV	1×10^{-16}	1×10^{-16}–1×10^{-6}	20 min	Present work

ranging from 0% to 85%. It was observed that the R_{ct} increases after hybridization with the PCR-amplified positive real samples, having elevated levels of transcript, whereas negligible change in the value is obtained with the negative patient samples (Fig. 7(C)). The percentage change in the R_{ct} value of the positive as well as negative patient samples in comparison to the R_{ct} value obtained for the target DNA indicates that the R_{ct} value for the positive patient samples has less percentage of variation in comparison to the target DNA while percentage of variation in R_{ct} value is more for negative patient samples (Table S1). Hence, the results obtained for the genosensor are comparable with the PCR-amplified real samples, which reveal that the present biosensing assay is suitable for the detection of BCR-ABL transcript. The comparison of response characteristics of the fabricated biosensor for CML detection have been done with other PANI based biosensors reported in the literature (Table 2), and it can be deduce that the fabricated biosensor shows better sensitivity and stability compared to other biosensors.

The potential of fabricated pDNA/t-PANI-NC/ITO bioelectrode relating to hybridization detection in human serum samples spiked with CML genomic DNA was investigated by using EIS (Fig. S8(A)). For the preparation of spiked samples, different concentrations of genomic DNA (10^{-15} M, 10^{-12} M, 10^{-8} M) in 10% human serum samples were prepared. After incubating the spiked DNA with pDNA/t-PANI-NC/ITO bioelectrode, the R_{ct} values was calculated for the different concentration. The increased values of the R_{ct} and reduced detection limit (10^{-15} M) for the serum samples in comparison to the R_{ct} values and detection limit (10^{-16} M) obtained for synthetic DNA samples is due to interference produced by different metabolites present in serum samples [48]. The curve obtained between the logarithmic values of various concentrations of the genomic DNA, spiked into diluted samples of human serum exhibited linear relationship with respect to their corresponding R_{ct} values. The results indicate that the recoveries were 93.6% (10^{-15} M), 94.8% (10^{-12} M) and 98.2% (10^{-8} M), suggesting the good accuracy of the developed method.

3.9. Reusability, reproducibility and stability of pDNA/t-PANI-NT/ITO

Regeneration studies of the prepared bioelectrodes were conducted by immersing it into the buffer solution containing 10 mM Tris-HCl and 1 mM of EDTA (pH 8.0) at 100 °C for 5 min. Then the fabricated bioelectrode was cooled down in the ice bath for ~ 30 min, the procedure removes the hybridized DNA *via* thermal denaturation. Hybridization after the regeneration process on the surface of bioelectrode was performed for six consecutive measurements and resultant change in R_{ct} value was recorded (Fig. S8(B)). The ΔR_{ct} value of the bioelectrode was found to decrease after each regeneration cycle and a reduction to about 89% of initial value after sixth cycle was observed. Although, a slight increase in the R_{ct} value for pDNA/t-PANI-NT/ITO bioelectrode after each regeneration cycle and decrease in the R_{ct} value after the hybridization indicates the incomplete denaturation and nonspecific adsorption of the target DNA on the surface of bioelectrode during the process. For determining the reproducibility of the fabricated pDNA/t-PANI-NT/ITO and pDNA/c-PANI-NT/ITO bioelectrodes, the relative standard deviation (RSD) was calculated. For RSD calculation two different concentrations of the cDNA (10^{-15} M and 10^{-11} M) were immobilized on five equally prepared bioelectrodes and their R_{ct} value was calculated. For pDNA/t-PANI-NT/ITO bioelectrode the values of RSD were found to be 5.7% and 4.8% for 10^{-15} M and 10^{-12} M concentration of cDNA, respectively and for pDNA/c-PANI-NT/ITO bioelectrode the values of RSD were found to be 7.6% and 6.4% for 10^{-15} M and 10^{-11} M concentration of cDNA respectively. The low RSD value for pDNA/t-PANI-NT/ITO bioelectrode in comparison to the pDNA/c-PANI-NT/ITO bioelectrode suggested the better repeatability of the proposed biosensor.

Further, the storage stability of the fabricated pDNA/t-PANI-NT/ITO and pDNA/c-PANI-NT/ITO bioelectrodes was investigated *via* EIS by recording five measurements each week for 40 days [42]. It was observed that the decrease in the signal response to 5% of its original R_{ct} value was found after 20days for pDNA/t-PANI-NT/ITO bioelectrode (Fig. 7(D); curve (i)) and 15days for pDNA/c-PANI-NT/ITO bioelectrode (Fig. 7(D); curve (ii)), when the electrode was stored in the refrigerator at 4 °C.

4. Conclusions

In summary, we have synthesized PANI-NT by using MnO_2 as a sacrificial template. The t-PANI NT were successfully electrophoretically deposited on ITO coated glass substrates. The morphological and electrochemical studies clearly reveal that the t-PANI NT have conducive morphology, improved conductivity, and better redox activity compared to PANI-MnO_2 NT and c-PANI-NT. Further t-PANI-NT/ITO electrodes were utilized as an impedimetric nucleic acid sensor for specific detection of CML. The fabricated genosensor was found to detect a wide range of complementary DNA concentration (10^{-6} to 10^{-16} M) with lower detection limit (10^{-16} M), good reproducibility and high stability (40 days). The biosensor was able to distinguish the clinical samples of CML positive and negative patients. Besides this, attempts are being made for the construction of highly sensitive genosensor with a wide variety of nucleic acid detection applications that may have importance in point-of-care diagnostics.

Acknowledgements

We thank Director, National Physical Laboratory, India for providing facilities. A.S. is thankful to CSIR, India, for the award of SRF. C.M. P. acknowledges the financial support received from Department of Science and Technology, India under DST-INSPIRE Faculty Project (Grant No. DST/INSPIRE/04/2015/000932). S.S. is grateful to UGC, India for the award of SRF. We thank Prof. B.D Malhotra (Delhi Technological University, Delhi, India) for useful discussions. Author thanks, Dr. Birendra Kumar Yadav (Rajiv Gandhi Cancer Institute and Research Centre, New Delhi) for providing the clinical patient samples.

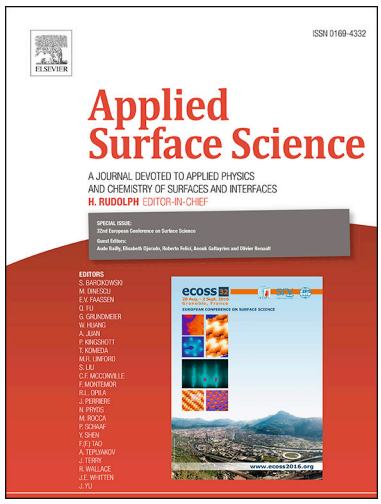
Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2018.05.038.

References

[1] D. Li, J. Huang, R.B. Kaner, Acc. Chem. Res. 42 (2008) 135–145.
[2] G.-R. Li, Z.-P. Feng, J.-H. Zhong, Z.-L. Wang, Y.-X. Tong, Macromolecules 43 (2010) 2178–2183.
[3] H.-W. Park, T. Kim, J. Huh, M. Kang, J.E. Lee, H. Yoon, ACS Nano 6 (2012) 7624–7633.
[4] W. Chen, R.B. Rakhi, H.N. Alshareef, J. Mater. Chem. A 1 (2013) 3315–3324.
[5] X. Feng, Y. Zhang, Z. Yan, N. Chen, Y. Ma, X. Liu, X. Yang, W. Hou, J. Mater. Chem. A 1 (2013) 9775–9780.
[6] V. Talwar, O. Singh, R.C. Singh, Sens. Actuators B: Chem. 191 (2014) 276–282.
[7] X. Shi, A.L. Briseno, R.J. Sanedrin, F. Zhou, Macromolecules 36 (2003) 4093–4098.
[8] Z. Zhang, J. Sui, L. Zhang, M. Wan, Y. Wei, L. Yu, Adv. Mater. 17 (2005) 2854–2857.
[9] L. Xia, Z. Wei, M. Wan, J. Colloid Interface Sci. 341 (2010) 1–11.
[10] L. Pan, H. Qiu, C. Dou, Y. Li, L. Pu, J. Xu, Y. Shi, Int. J. Mol. Sci. 11 (2010) 2636–2657.
[11] W. Chen, R.B. Rakhi, H.N. Alshareef, J. Mater. Chem. C 117 (2013) 15009–15019.
[12] I. Tiwari, M. Gupta, C.M. Pandey, V. Mishra, Dalton Trans. 44 (2015) 15557–15566.
[13] C. Dhand, M. Das, G. Sumana, A.K. Srivastava, M.K. Pandey, C.G. Kim, M. Datta, B.D. Malhotra, Nanoscale 2 (2010) 747–754.
[14] Y. Zhu, J. Zhang, Y. Zheng, Z. Huang, L. Feng, L. Jiang, Adv. Funct. Mater. 16 (2006) 568–574.
[15] L. Li, Y. Shi, L. Pan, Y. Shi, G. Yu, J. Mater. Chem. B 3 (2015) 2920.
[16] C. Dhand, M. Das, M. Datta, B.D. Malhotra, Biosens. Bioelectron. 26 (2011) 2811–2821.
[17] T. Sen, S. Mishra, N.G. Shimpi, RSC Adv. 6 (2016) 42196–42222.
[18] Y.-Z. Long, M.-M. Li, C. Gu, M. Wan, J.-L. Duvail, Z. Liu, Z. Fan, Prog. Polym. Sci. 36 (2011) 1415–1442.
[19] R. Liu, S.B. Lee, J. Am. Chem. Soc. 130 (2008) 2942–2943.
[20] M. Yang, J. Ma, C. Zhang, Z. Yang, Y. Lu, Angew. Chem. Int. Ed. 44 (2005) 6727–6730.
[21] M.W.N. Deininger, J.M. Goldman, J.V. Melo, Blood 96 (2000) 3343–3356.
[22] J.P. Radich, T. Gooley, E. Bryant, T. Chauncey, R. Clift, L. Beppu, S. Edmands, M.E.D. Flowers, K. Kerkof, R. Nelson, Blood 98 (2001) 1701–1707.
[23] J. Chen, J. Zhang, L. Huang, X. Lin, G. Chen, Biosens. Bioelectron. 24 (2008) 349–355.
[24] A.S. Advani, A.M. Pendergast, Leuk. Res. 26 (2002) 713–720.
[25] A.S. Corbin, A. Agarwal, M. Loriaux, J. Cortes, M.W. Deininger, B.J. Druker, J. Clin. Investig. 121 (2011) 396.
[26] M.A. Tainsky, Biochim. Biophys. Acta 1796 (2009) 176–193.
[27] M. Shamsipur, V. Nasirian, A. Barati, K. Mansouri, A. Vaisi-Raygani, S. Kashanian, Anal. Chim. Acta 966 (2017) 62–70.
[28] D. Ozkan-Ariksoysal, Y.U. Kayran, F.F. Yilmaz, A.A. Ciucu, I.G. David, V. David, M. Hosgor-Limoncu, M. Ozsoz, Talanta 166 (2017) 27–35.
[29] T. Yu, H. Zhang, Z. Huang, Z. Luo, N. Huang, S. Ding, W. Feng, Electroanalysis 29 (2017) 828–834.
[30] Y.-X. Chen, K.-J. Huang, F. Lin, L.-X. Fang, Talanta 175 (2017) 168–176.
[31] C.M. Pandey, G. Sumana, K.N. Sood, B.D. Malhotra, Thin Solid Films 519 (2010) 1178–1183.
[32] A.S. Ghrera, C.M. Pandey, B.D. Malhotra, Sens. Actuators B: Chem. 266 (2018) 329–336.
[33] K.Y.P.S. Avelino, I.A.M. Frias, N. Lucena-Silva, R.G. Gomes, C.P. de Melo, M.D.L. Oliveira, C.s.A.S. Andrade, Colloids Surf. B 148 (2016) 576–584.
[34] R. Singh, R. Verma, G. Sumana, A.K. Srivastava, S. Sood, R.K. Gupta, B.D. Malhotra, Bioelectrochemistry 86 (2012) 30–37.
[35] N. Hui, X. Sun, S. Niu, X. Luo, ACS Appl. Mater. Interfaces 9 (2017) 2914–2923.
[36] C. Dhand, P.R. Solanki, M.K. Pandey, M. Datta, B.D. Malhotra, Electrophoresis 31 (2010) 3754–3762.
[37] L. Zhang, M. Wan, Nanotechnology 13 (2002) 750.
[38] R. Singh, R. Prasad, G. Sumana, K. Arora, S. Sood, R.K. Gupta, B.D. Malhotra, Biosens. Bioelectron. 24 (2009) 2232–2238.
[39] A. Sumboja, C.Y. Foo, J. Yan, C. Yan, R.K. Gupta, P.S. Lee, J. Mater. Chem. 22 (2012) 23921–23928.
[40] J. Stejskal, J. Prokes, M. Trchovs, Polym. Degrad. Stab. 102 (2014) 67–73.
[41] S. Solanki, C.M. Pandey, A. Soni, G. Sumana, A.M. Biradar, Microchim. Acta 183 (2015) 167–176.
[42] A. Soni, C.M. Pandey, S. Solanki, G. Sumana, RSC Adv. 5 (2015) 45767–45774.
[43] B.G. Choi, H. Park, M.H. Yang, Y.M. Jung, S.Y. Lee, W.H. Hong, T.J. Park, Nanoscale 2 (2010) 2692–2697.
[44] M.U.A. Prathap, R. Srivastava, B. Satpati, Electrochim. Acta 114 (2013) 285–295.
[45] Y. Bo, H. Yang, Y. Hu, T. Yao, S. Huang, Electrochim. Acta 56 (2011) 2676–2681.
[46] H.-L. Shuai, K.-J. Huang, Y.-X. Chen, J. Mater. Chem. B 4 (2016) 1186–1196.
[47] C.M. Pandey, I. Tiwari, G. Sumana, RSC Adv. 4 (2014) 31047–31055.
[48] N. Prabhakar, H. Singh, B.D. Malhotra, Electrochem. Commun. 10 (2008) 821–826.
[49] Y. Hao, B. Zhou, F. Wang, J. Li, L. Deng, Y.-N. Liu, Biosens. Bioelectron. 52 (2014) 422–426.
[50] Q. Zhu, F. Gao, Y. Yang, B. Zhang, W. Wang, Z. Hu, Q. Wang, Sens. Actuators B: Chem. 207 (2015) 819–826.
[51] J. Wilson, S. Radhakrishnan, C. Sumathi, V. Dharuman, Sens. Actuators B: Chem. 171 (2012) 216–222.
[52] Q. Zheng, H. Wu, Z. Shen, W. Gao, Y. Yu, Y. Ma, W. Guang, Q. Guo, R. Yan, J. Wang, Analyst 140 (2015) 6660–6670.

Accepted Manuscript



Full Length Article

Electron irradiation induced wall-to-wall joining of multiwalled carbon nanotubes

Reetu Kumari, Pawan K. Tyagi, Nitin K. Puri

PII: S0169-4332(18)31331-X
DOI: https://doi.org/10.1016/j.apsusc.2018.05.055
Reference: APSUSC 39331

To appear in: *Applied Surface Science*

Received Date: 1 March 2018
Revised Date: 27 April 2018
Accepted Date: 8 May 2018

Please cite this article as: R. Kumari, P.K. Tyagi, N.K. Puri, Electron irradiation induced wall-to-wall joining of multiwalled carbon nanotubes, *Applied Surface Science* (2018), doi: https://doi.org/10.1016/j.apsusc.2018.05.055

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Electron irradiation induced wall-to-wall joining of multiwalled carbon nanotubes

Reetu Kumari, Pawan K. Tyagi[#], Nitin K. Puri

Department of Applied Physics, Delhi Technological University, Delhi, 110042

Abstract:

In present work, wall-to-wall joining of two parallel multiwalled carbon nanotubes (MWCNTs) has been successfully demonstrated using electron beam inside high resolution transmission electron microscope. The used MWCNTs have been grown using cobaltocene and benzene as catalyst and carbon precursor, respectively via thermal CVD technique. We have proposed that in present case, irradiation induced joining is achieved if adequate balance between sputtered carbon atoms from the outer walls of MWCNTs and generation as well as migration of defects along the tube axis is obtained. Casmir effect is not found to play dominating role in the joining process. The proposed method can potentially be used for developing wall-to-wall interconnection between MWCNTs using electron beam.

***Keywords*:** Carbon nanotubes, electron-irradiation, wall-to-wall interconnection, *in-situ* transmission electron microscopy

[#] Email: pawantyagi@dtu.ac.in, pawankumartyagi@gmail.com

ACCEPTED MANUSCRIPT

1. Introduction:

The interconnected carbon nanotubes (CNTs) are considered to be building blocks of many nanoscale electronic [1.2] and mechanical devices [3]. This has triggered the research interest in cross-linking, fusion, welding and joining of CNTs. Previously, it is reported that CNTs can be interconnect by chemical functionalization [4], heat welding [5,6], laser beam [7] atomic force microscope [8], ion [9,10] and electron irradiation [11,12,13]. In order to further expand the applications of multiwalled carbon nanotubes (MWCNTs), key aspect will be the post-growth modification techniques which can join outer walls of MWCNT and can for the wall-to-wall junction in a controlled manner. Recently, many interesting phenomena such as thinning, breaking, shrinkage [14], cutting [15] and welding [12] have been demonstrated by using electron beam inside high resolution transmission electron microscope (HRTEM). These irradiation-induced structural modifications are accomplished by production of defects such as point defects, interstitial-vacancy pair, saturation of dangling bonds, generation as well as migration of pentagon-heptagon pairs and evaporation of carbon atoms. In order to interconnect the MWCNTs, joining has great importance as they can be used as interconnects in many electronic devices. As mentioned-above, in joining phenomenon, creation of defects, control over damage and saturation of dangling bonds play a fundamental role. Irradiation-induced dangling bonds in carbon nanostructures have inherent ability to form the structures with carbon atoms of different hybridizations. Such reconstruction of bonds by C atom is only achieved thorough annealing which leads the possibility of cross-linking of carbon nanotubes using electron beam at high temperature [16]. As reported earlier, cross-linking of two CNTs under e-beam irradiation retaining their graphitic structure can only be achieved at high temperature [16,17,18]. This is due to the stability of nanotubes under e-beam at high temperature which is governed by generation and annealing of vacancy-interstitial pair [18]. Previously, joining of CNTs in different geometry

ACCEPTED MANUSCRIPT

such as: cap-to-cap and cap-to-wall has been demonstrated by Chuanhong Jin *et al*.[19]. They have reported that generation and migration rate of pentagon-heptagon pairs along the CNT axis should be higher than that of evaporation rate of carbon diamer from the tubes. Recently, it has been reported that annealing of vacancy-interstitial pairs is only possible if specimen temperature is above 300 ℃ [20]. Yoon *et al.* have successfully demonstrated the zipper type joining of two separate SWCNTs using electron beam irradiation at high temperature [11]. Terrones *et al.* also simulated and experimentally demonstrated the fusion of CNT under electron beam inside HRTEM operated at 1.25 MeV keeping specimen temperature 800 ℃ [17]. As reported earlier, at higher temperature interstitial carbon atoms are more mobile and can repair defects more quickly [21]. This means that at higher temperature CNTs not only maintain their stability but also exhibit reconstruction and minimize irradiation induced amorphization. Evora *et al.* reported the functionalization carbon nanofibers with incorporation of oxygen at the surface by oxidizing them in air at 350°C under the influence of a 3 MeV electron beam [22]. At such high energy irradiation, some modifications such as cutting, bending or fusion have been observed at discrete points through functionalization which was highly localized and did not lead to modifications on the bulk nanofiber structure. Recently, electron irradiation on Fe_3C filled-MWCNT has been performed inside transmission electron microscope by Singh et.al. [23]. They have demonstrated the buckling, thinning, bending and breaking of nanotube as well as Ostwald ripening in encapsulated nanorod. In this report an individual MWCNT has been irradiated and focus was to study the response of encapsulated nanorod under electron irradiation. In other reports, crosslinking, bridging or functionalization between two or more SWCNTs has been demonstrated [11,17,19]. However, bridging or crosslinking can provide a very small area of contact on the junction point or could be limited to few lattice points. In some cases, soldering is achieved through deposition of amorphous carbon [12] or through metal particles [24]. In the

ACCEPTED MANUSCRIPT

present case, in demonstrated sidewall joining, the graphitic structures of walls of individual CNTs are preserved and most importantly the junction area on side walls is much higher than crosslinking. As irradiation proceeds joining propagates along the side walls of nearby CNTs. Apart for e-beam, structural modifications and crosslinking or bridging between CNTs under ion bean irradiation have also been reported. Zhang *et al*. reviewed the recent progress in the understanding of ion-irradiation-inducing phenomena in CNTs [25]. In this review, inter-tube linking through dangling bond at distorted lattice [26,27], welding of CNTs crossing each other and functionalization [28,29] have been reviewed. Such linking or welding is explained mainly for SWCNTs which affects only the region where they cross each other unlike side-wall joining and many of them adopted computational methods [25,26,27]. Apart from these, formation of crossing and parallel junction between MWCNTs by ion beam irradiation at elevated temperature has been reported [30,31]. However, in their report it was observed that below 600 K, the irradiated MWCNTs were greatly destroyed even at low irradiation dose. At 800 K it was improved but the graphitic walls seemed significantly distorted. Moreover, ion beam irradiation process needs a complicated laboratorial setups and high dose of ions requires much longer duration even many hours. Long exposure to ion beam causes amorphization or extensive distortions in the graphitic structure of CNTs. For optimum benefits of such CNT-junctions in applications, it is necessary to preserve the graphitic structure of CNTs to the maximum possible extent which can be achieved in e-beam irradiation only in few seconds and comparatively at lower temperature.

Side wall linking of CNTs can be potentially exploited in recent advanced applications. Barra *et al*. have investigated the effect of the addition of CNTs with Polyhedral oligomeric silsesquioxane to toughen the epoxy adhesives [32]. The addition of CNTs was found to be responsible for effective bridging function across the fracture surface of the bonded joints and enhancing the mechanical properties. The joined or crosslinked CNTs can

ACCEPTED MANUSCRIPT

improve these properties providing more junctions and boost mechanical strength as well. In another recent work, effect of MWCNTs on the photooxidative behavior of nanocomposites based on the polymorph of isotactic polypropylene has been studied. Due to presence of MWCNTs, the polymer shows high stability to the photo-oxidative degradation and show electrical conductivity [33]. Therefore, increasing contact regions between CNTs can enhance the electrical properties. Guadagno *et al.* have designed new epoxy by mixing the Tetra-Glycidylmethylenedianiline with the epoxy monomer 1–4 butanedioldiglycidyl ether and 4,4′-diaminodiphenyl sulfone was used as hardener agent. In this epoxy resin, CNTs and other carbon nanostructures were mixed to obtain an enhancement in relevant physical properties. They reported that after addition of CNTs, moisture content by entrapping parts of resin reduced. Such entrapping results in inhibition of the interaction of the reactive sites of the resin with the water molecules [34]. Junctions of CNTs may cause an increase in the more extended zones in contact with CNTs and thus more epoxy networks. Therefore, in such recent CNT based composite related applications, junctions limit the CNT to flow inside the polymeric matrix. This can give positive benefit on the formation of aggregate and improve dispersions or the damping characteristics reducing the water sorption as well as electrical and mechanical properties of such nanocomposites which could perform as piezoresistive or sensors.

It is clear from above discussions, in order to further enhance the applications of MWCNTs, wall-to-wall joining between MWCNTs has great importance and a simple way is desirable for joining them in such a manner that they retain their crystallinity. To the best of our knowledge, e-beam irradiation induced wall-to-wall joining of MWCNTs retaining crystalline structure has not been demonstrated till date. Therefore, the prime focus of present work is to demonstrate the wall-to-wall joining of two parallel by using electron beam in controlled manner. In the reported work, joining of two MWCNTs is demonstrated inside

ACCEPTED MANUSCRIPT

HRTEM operated at 200 kV and electron beam of energy 200 keV was used. The irradiation induced drifting and joining of the MWCNTs were comprehensively monitored in real time.

.

2. Experimental Details:

The used unfilled/Co-filled MWCNTs were grown in modified thermal chemical vapour deposition (Thermal CVD) system. The specifications of used experimental set up have been described previously [35]. A mixture of cobaltocene and benzene in the ratio of 0.015 g/ml was taken as the precursor. The ultra-sonicated mixture was first introduced into the pre-heating zone (150–200 °C) for evaporation. The vapors were then carried into the hot zone (850 °C) using carrier gases Ar/H_2 (100/50 sccm). After completion of the process, the reactor was allowed to cool down to room temperature in the presence of Ar gas. The deposited black powder was scratched from the walls of the quartz tube. A small amount of the scratched sample was dispersed in isopropyl alcohol and then sonicated for 10 minutes for TEM measurement. After sonication one drop of the dispersion was casted onto carbon-coated copper grid of 300 mesh. After evaporating IPA, grid was transferred to the specimen holder for electron irradiation and microscopy measurement. The black powder was confirmed as MWCNTs in microscopy measurements. In-situ electron irradiation induced wall-to-wall joining of MWCNTs as well as imaging was carried out in high-resolution transmission electron microscope (JEOL JEM 2010 operated at 200 kV). In present case, electron beam of energy 200 keV was used and irradiation was carried out at high temperature (400 °C) and beam current density was approximately 120-140 pA/cm^2. The current density is influenced by parameters like, filament power, spot-size (C1), magnification, and condenser aperture. Such details are displayed in the system. In the present case 110-113 pA was the operating beam current and condenser aperture was 120 µm which corresponds to this value of current density. The beam diameter was in micron size and

ACCEPTED MANUSCRIPT

the diameter of beam spot on sample was nearly 150 nm. Therefore, the whole area displayed in the TEM images are beam affected.

3. Results and discussion

Wall-to-wall joining of two adjacent MWCNTs separated at a distance nearly 1.5 nm are demonstrated in Figure 1. The joining and reference points are 4 nm apart and marked as 1 and 2 in Fig. 1(a), respectively. Before irradiation, wall-to-wall distance of two parallel MWCNTs were found to be ~1.5 nm which is larger than the Van der Waals distance of 3.4 Å and much smaller than the separation required for Casmir effect to dominate which is tens of nanometer [36]. As irradiation proceeded, initially these MWCNTs came closer and joined together to form junction between the outer walls at point 1. As irradiation further proceeded, junction length increased as marked in Fig. 1c, MWCNT 1 and 2 were 6 and 10 walled, respectively. After continuous irradiation for 240 seconds, at junction point 1 total number walls were found to decrease from 16 to15 as shown in Figs. 1c&f. Therefore, it can be presumed that the outer walls of both MWCNTs were merged. After 480 seconds, MWCNTs were completely joined in irradiated area as shown in Figs. 1d&e. High resolution image of joining point clearly revealed that outer walls of MWCNTs were merged and resulted in formation of edge dislocation.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

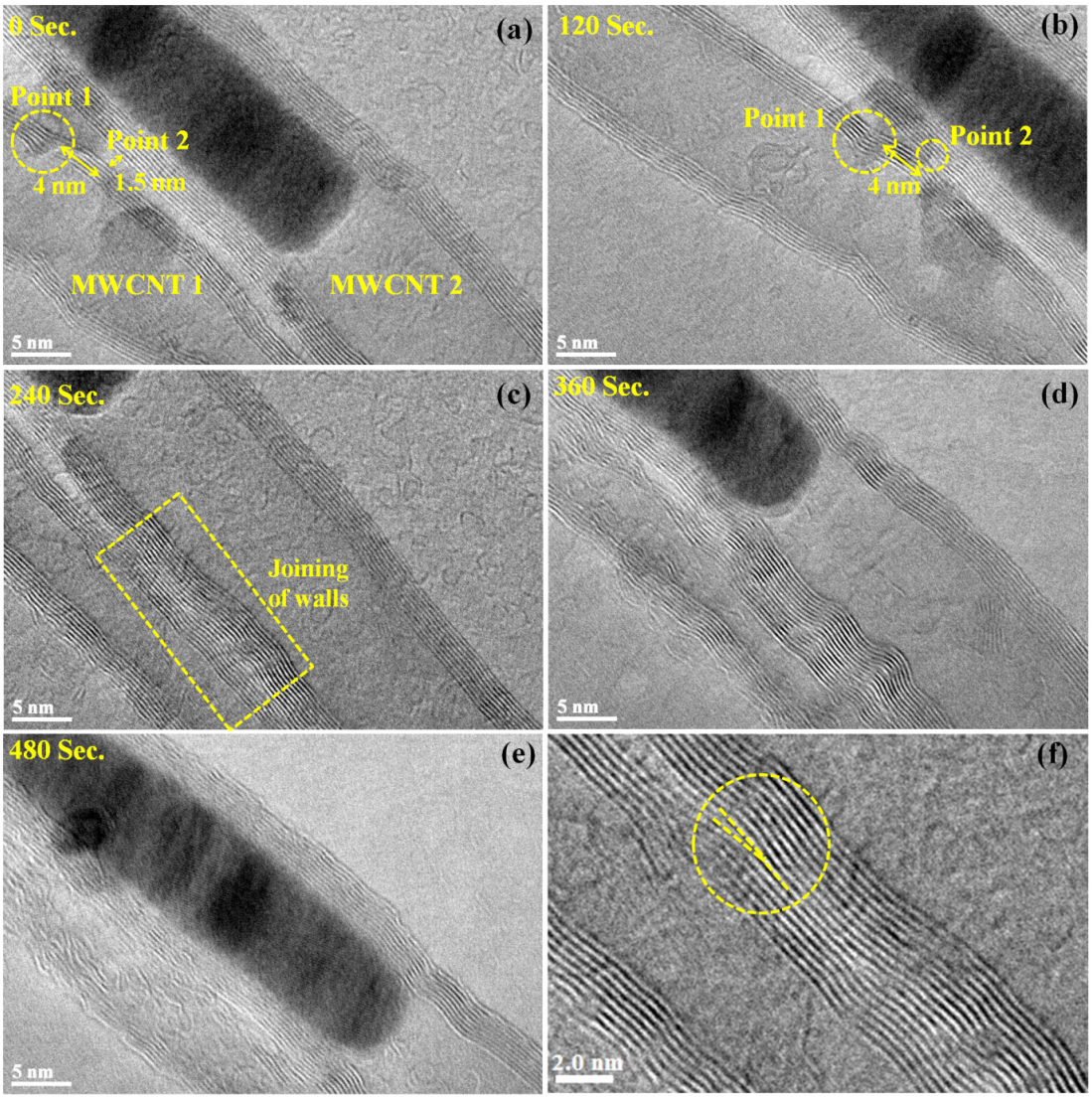


Figure 1. (a) TEM image of two MWCNTs separated at 1.5 nm distance, (b) after 2 min. MWCNTs touched at a point highlighted with a circle, (c) after 4 min., outer walls joined together. Image captured after (d) 6 min. and (e) 8 min. (f) high resolution image depicts the wall-to-wall joining.

In order to validate the joining phenomenon, the study was further extended to another pair of tubes separated at a distance of 2-5 nm. The video of *in-situ* electron irradiation induced joining phenomena was captured for 21 minutes. Images captured at time interval of 2 minutes have been shown in Figure 2. As evident from Fig 2, MWCNTs vibrated and moved towards each other. This might have happened due to the transfer of

ACCEPTED MANUSCRIPT

kinetic energy from e-beam to MWCNTs. Here, it is noteworthy to mention that separation between these MWCNTs is not constant and gradually increased from upper part to lower part of the image. Similar to the observation in Fig 2, joining was started at point where effective separation between MWCNTs is nearly 2 nm.

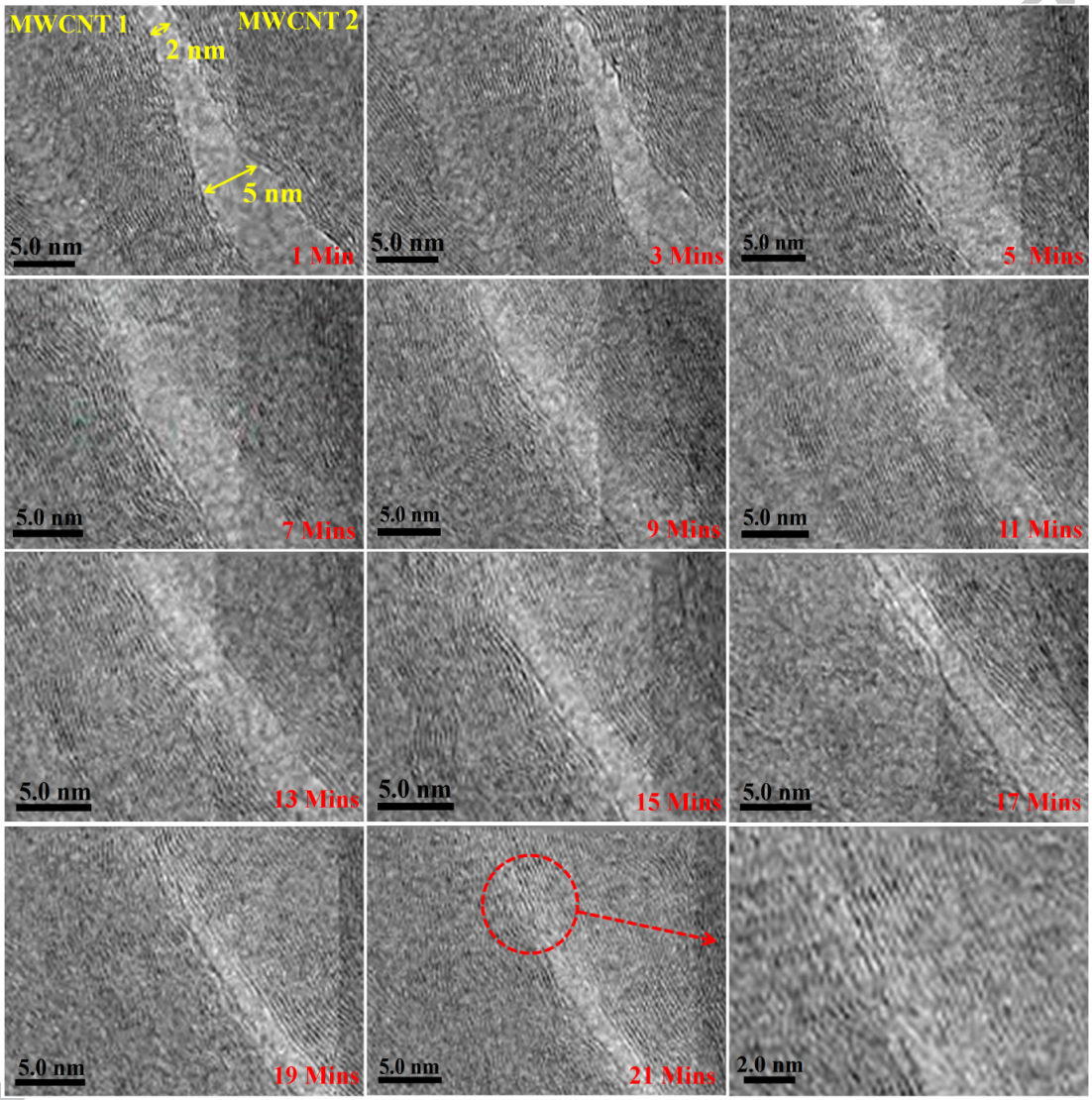


Figure 2. High resolution images of joining process captured at various time intervals from the real-time movie which was recorded during sustained electron-irradiation at high temperature (400 °C). Irradiation period is mentioned in panels.

ACCEPTED MANUSCRIPT

Another example of wall-to-wall joining is demonstrated in Fig.3 where 2nm apart MWCNTs were joined together. As evident from Fig 3b, outer walls of MWCNTs were covalently linked and edge dislocation was also observed.

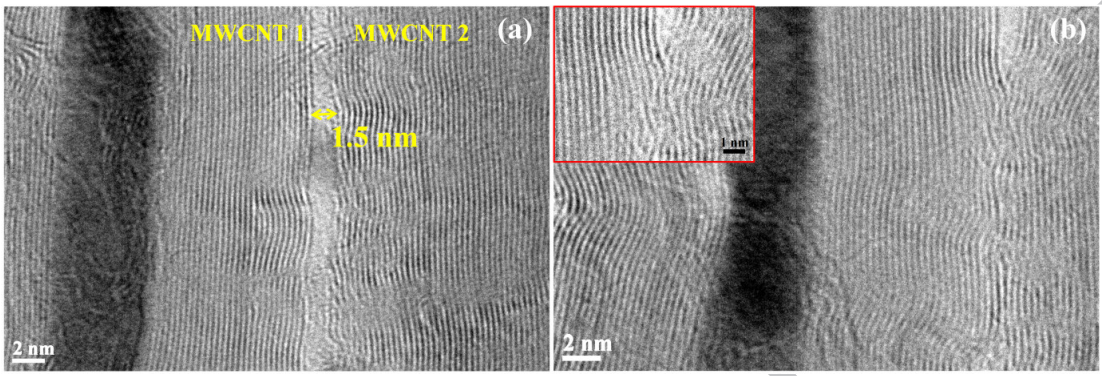
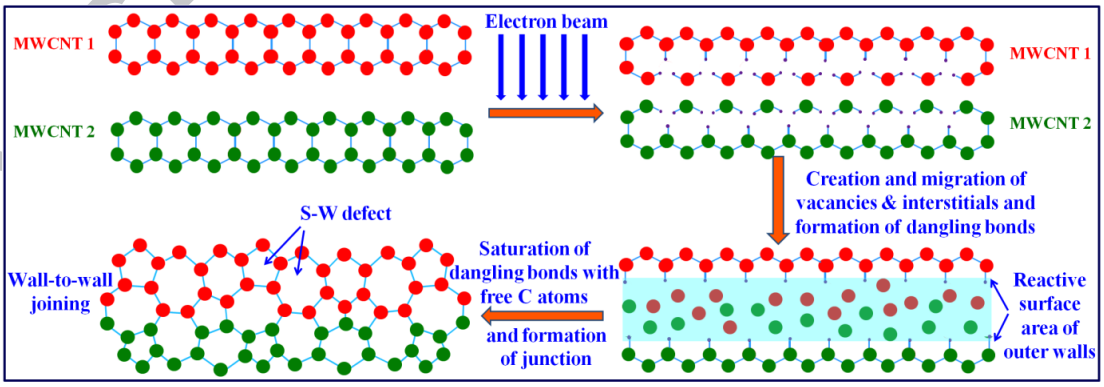


Figure 3. High resolution image of two parallel MWCNTs (a) before and (b) after joining. Magnified image of wall-to-wall junction is shown in the inset of (b).

In order to understand the accomplishment of joining process, there could be two possibilities; first is junction formation due to sputtering of carbon atoms from the outer wall of the tubes and generation as well as migration of defects such as pentagon-heptagon pairs and second is Casimir effect. If we consider that joining happened due to Casimir effect then in such case total number of walls would be constant and phenomena such as merging of outer walls and formation of edge dislocation should not be observed. Therefore, the possibility of Casimir force cannot be considered as dominating factor and can be ruled out. Explicit interpretation of joining process is not possible without the precise knowledge of the atomic defects produced during electron irradiation. Therefore, we need to understand mechanism and nature of defects produced in outer as well as inner walls of MWCNTs. Generally, irradiation induced defects are vacancies and interstitials. As irradiation proceeds interstitials move between adjacent graphitic lattice and also diffuse on the periphery of the MWCNTs. Vacancies and interstitials moving between the lattice planes will not contribute in wall-to-

ACCEPTED MANUSCRIPT

wall joining process. However, atoms diffusing on the periphery of CNT will make new bond structures. If diffusion of carbon atoms on periphery is taken into account then after joining number of walls would be increased. The reduction in the number of the walls reveals that joining process might have occurred due to the creation of the interstitials and sputtering of carbon atoms from outer walls which led to the cross linking between outer walls of two MWCNTs as demonstrated in Fig.4. As irradiation further proceeds at the point of joining, these cross-linked walls will merge in to a single wall. It has been reported by Chuanhong Zin *et al.* that in order to join CNTs in different geometry such as: cap-to-cap and cap-to-wall generation and migration rate of pentagon-heptagon pairs along the CNT should be higher than that of sputtering rate of carbon atoms [20]. In present case, during electron irradiation, sputtering rate of carbon atoms seem to be high and only small fraction of carbon atoms knock-out as interstitials. Eventually, generation and migration rate of pentagon-heptagon pairs along tube axis will be lowered. This type of loss of carbon atoms led to reconstruction of new coherent structure containing Stone-Wales (SW) defects. As a consequence, deformation and shrinkage of tube walls occurred as evident from Figs. 1c&d. As irradiation further proceeded, first, these defects and sputtered carbon atoms made cross-link with outer wall of adjacent MWCNT and then the outer walls of these two MWCNTs merged in to single wall followed by formation of edge dislocation.



ACCEPTED MANUSCRIPT

Figure 4. Schematic representation of steps involved in wall-to-wall joining of two parallel MWCNTs.

4. Conclusion

In this report we have successfully achieved the wall-to-wall joining of two parallel MWCNTs inside HRTEM. We propose that irradiation induced joining occurred due to elegant balance between sputtering of carbon atoms from the outer walls as well as generation and migration of the Stone-Wales defects along the tube axis rather than Casmir effect. This observation was evident from the decrement in total number of walls in the junction area. We envision that reported work might be a forwarding step in developing interconnections between MWCNTs using electron beam which could be of great benefit in the field of electronic interconnects.

Acknowledgments

The authors would like to thank Prof. P. V. Satyam and Mr. Puspendu Guha for carrying out TEM measurements at IOP Bhubaneswar.

References:

1. Z. Yao, H. W. C. Postma, L. Balents, C. Dekker, Carbon nanotube intramolecular junctions, Nature 402 (1999) 273-276.
2. L. Chico, V.H. Crespi, L.X. Benedict, S.G. Louie, M. L. Cohen, Pure carbon nanoscale devices: Nanotube heterojunctions, Phys Rev Lett. 76 (1996) 971-974.
3. Z. Lin, X. Gui, Q. Gan, W. Chen, X. Cheng, M. Liu, Y. Zhu, Y. Yang, A. Cao, Z. Tang, In-Situ Welding Nanotubes into a Porous Solid with Super-High Compressive Strength and Fatigue Resistance, Sci. Rep. 5 (2015) 11336.

ACCEPTED MANUSCRIPT

4. P. W. Chiu, G. S. Duesberg, U. Dettlaff-Weglikowska, S. Roth, Interconnection of carbon nanotubes by chemical functionalization, Appl. Phys. Lett. 80 (2002) 3811-13.

5. X. Yang, Z. Han, Y. Li, D. Chen, P. Zhang, A. To, Heat welding of non- orthogonal X-junction of single-walled carbon nanotubes. Phys. E 46 (2012) 30–32.

6. H. Hirayama, Y. Kawamoto, Y. Ohshima, K. Takayanagi, Nanospot welding of carbon nanotubes, Appl. Phys. Lett.79 (2001) 1169-1171.

7. Y. Yuan, J. Chen, Nano-Welding of Multi-Walled Carbon Nanotubes on Silicon and Silica Surface by Laser Irradiation, Nanomaterials 6 (2016) 1-7.

8. H.W.C. Postma, M. Jonge, Z. Yao, C. Dekker, C. Electrical transport through carbon nanotube junctions created by mechanical manipulation. Phys. Rev. B 62 (2010) R10653–R10656.

9. A. V. Krasheninnikov, K. Nordlund, and J. Keinonen, F. Banhart, Ion-irradiation-induced welding of carbon nanotubes, Phys. Rev. B 66 (2002) 245403-1-6.

10. R. Makala, G. G. Pethuraja, J. D. 'arcy-Gall, I. Petro, Nanomachining carbon nanotubes with ion beams, App. Phy. Lett. 84 (2004) 4484–4486.

11. M. Yoon, S. Han, G. Kim, S. B. Lee, S. Berber, E. Osawa, Jisoon Ihm, M. Terrones, F. Banhart, J. C. Charlier, N. Grobert, H. Terrones, P. M. Ajayan, D. Tománek, Zipper Mechanism of Nanotube Fusion: Theory and Experiment, Phys Rev Lett. 92 (2004) 075504-1-4.

12. F. Banhart. The Formation of a Connection between Carbon Nanotubes in an Electron Beam, Nanoletters 1 (2002) 329-332.

13. B. Li, Y. Feng, Ke-wang Ding, G. Qian, Xue-bin Zhang, Yan-fang Liu, Effect of electron beam irradiation on multi-walled carbon nanotubes, Trans. Nonferrous Met. Soc. China 24 (2014) 764-769.

ACCEPTED MANUSCRIPT

14. A. Singh, R. Kumari, V. Kumar, L. Krishnia, Z. Naqvi, A.K. Panwar, U.M. Bhatta, A. Ghosh, P.V. Satyam, P. K. Tyagi, Electron irradiation induced buckling, morphological transformation, and inverse Ostwald ripening in nanorod filled inside carbon nanotube, App. Sur. Sci. 360 (2016) 1003-1008

15. T. D. Yuzvinsky, A. M. Fennimore, W. Mickelson, C. Esquivias, and A. Zett, Precision cutting of nanotubes with a low-energy electron beam, Appl. Phys. Lett. 86 (2005) 053109.

16. A. V. Krasheninnikov, F. Banhart, Engineering of nanostructured carbon, materials with electron or ion beams, Nature Materials 6 (2007) 723-33.

17. M. Terrones, F. Banhart, N. Grobert, J.-C. Charlier and P.M. Ajayan, Molecular Junctions by Joining Single-Walled Carbon Nanotubes, Phys. Rev. Lett. 89 (2002) 075505-1-4.

18. A.V. Krasheninnikov, K. Nordlund, Ion and electron irradiation-induced effects in nanostructured materials, J. App. Phys. 107 (2016) 071301-70.

19. C. Jin, K. Suenaga, S.Iijima, plumbing carbon nanotubes, Nature Nanotechnology 3 (2008) 17-21.

20. F. Banhart, Irradiation effects in carbon nanostructures, Rep. Prog. Phys. 62 (1999) 1181–1221.

21. G.S. Roberts, P. Singjai, Joining carbon nanotubes, Nanoscale 3 (2011) 4503-4514.

22. M.C. Evora, D. Klosterman, K. Lafdi, L. Li, J.L. Abot, Functionalization of carbon nanofibers through electron beam irradiation, Carbon 48 (2010) 2037-46.

23. A. Singh, R. Kumari, V. Kumar, L. Krishnia, Z. Naqvi, A.K. Panwar, U.M. Bhatta, A. Ghosh, P.V. Satyam, P. K. Tyagi, Electron irradiation induced buckling, morphological transformation, and inverse Ostwald ripening in nanorod filled inside carbon nanotube, App. Sur. Sci. 360 (2016) 1003-1008.

ACCEPTED MANUSCRIPT

24. J. A. Rodriguez-Manzo, M. S. Wang, F. Banhart, Y. Bando, D. Golberg, Multibranched junctions of carbon nanotubes via cobalt particles, Adv. Mater. 21 (2009) 4477.

25. Y. Zhang, L. Chen, Z. Xu, Y. Li, M. Shan, L. Liu, Q. Guo, G. Chen, Z. Wang, Ch. Wang, Effects of ion irradiation on carbon nanotubes: a review, Int. J. Mat. Prod. Tech. 45 (2012) 1-30.

26. E. Salonen, A.V. Krasheninnikov, K. Nordlund, Ion-irradiation-induced defects in bundles of carbon nanotubes, Nuclear Instruments and Methods in Physics Research B 193 (2002) 603–608.

27. Sammalkorpi, M., Krasheninnikov, A.V., Kuronen, A., Nordlund, K. and Kaski, K. Irradiation-induced stiffening of carbon nanotube bundles', Nuc. Inst. Meth. Phy. Res. B, 228 (2005) 142–45.

28. A. V. Krasheninnikov, K. Nordlund, and J. Keinonen, Ion-irradiation-induced welding of carbon nanotubes, Phy Rev. B, 66 (2002) 245403,

29. Ni, B. R. Andrews, D. Jacques, D. Qian, M.B.J. Wijesundara, Y.S. Choi, L. Hanley, S.B. Sinnott, A combined computational and experimental study of ion-beam modification of carbon nanotube bundles, J. Phys. Chem. B 105 (2001) 12719–25.

30. Z. Ni, Q. Li, L. Yan, J. Gong, D. Zhu, Welding of multi walled carbon nanotubes by ion beam irradiation, Carbon 46 (2008) 376-378.

31. A. Ishaq, Z.C Ni, L. Yan, L.J. Gong, D.Z. Zhu, Constructing carbon nanotube junctions by Ar ion beam irradiation, Rad. Phys. Chem. 79 (2010) 687–691.

32. G. Barra, L. Vertuccio, U. Vietri, C. Naddeo, H. Hadavinia, L. Guadagno, Toughening of epoxy adhesives by combined interaction of carbon nanotubes and silsesquioxanes, Materials 10 (2017) 1131.

33. C. Naddeo, L. Vertuccio, G. Barra, L. Guadagno, Nano-charged polypropylene application:realistic perspectives for enhancing durability, materials 10 (2017) 943.

ACCEPTED MANUSCRIPT

34. L. Guadagno, C. Naddeo, M. Raimondo, G. Barra, L. Vertuccio, S. Russo, K. Lafdi, V Tucci, G. Spinelli, P. Lamberti, Nanotechnology 28 (2017) 094001.

35. R. Kumari L. Krishnia, V. Kumar, S. Singh, H. K. Singh, R. K. Kotnala, R. R. Juluri, U. M. Bhatta, P. V. Satyam, B.S. Yadav, Zainab Naqvi, P. K. Tyagi, Fe_3C-filled carbon nanotubes: permanent cylindrical nanomagnets possessing exotic magnetic properties, Nanoscale 8 (2016) 4299-4310

36. G. L. Klimchitskaya and V. M. Mostepanenko, van der Waals and Casimir interactions between two graphene sheets, Phys. Rev. B 87 (2013) 075439-1-7.

ACCEPTED MANUSCRIPT

Surface Topography: Metrology and Properties

ACCEPTED MANUSCRIPT

Experimental comparative study of chrome steel pin with and without chrome plated cast iron disc in situ fully flooded interface lubrication

To cite this article before publication: Roop Lal *et al* 2018 *Surf. Topogr.: Metrol. Prop.* in press https://doi.org/10.1088/2051-672X/aac60a

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2018 IOP Publishing Ltd**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.107 on 22/05/2018 at 08:16

Experimental comparative study of chrome steel pin with and without chrome plated cast iron disc in situ fully flooded interface lubrication

Roop Lal and R. C. Singh

Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

Corresponding author: rooplalrana@gmail.com (R. Lal).

Abstract

This paper reports the findings of an experimental study pertaining to the tribo dynamic investigations of contacts formed between chrome base pin and cast iron/chrome plated cast iron discs in situ conditions. Pin on disc geometrical configurations were employed for conducting the experiments at different operating parameters using SAE-15W40 lubricating oil. The fully flooded lubricating condition was insured by directed supply of lubricant at the interface. The friction coefficient, wear rate, vibration amplitude, temperature rise, and sound intensity were measured at different loads. (varying in the range of contact pressure 1.9 - 3.1 MPa) and sliding speeds (2-10 m/s) for the travelling distances between 1000- 5000 m. Based on the experimental investigations, reductions in the friction coefficient and wear rate were recorded for the chrome steel pin/chrome plated cast iron disc in comparison to chrome steel pin/cast iron disc contact. However, increase in the temperature rise, vibration, and noise were recorded with chrome steel pin/chrome plated cast iron disc contact.

Key Words: Friction, Wear, Pin-on-Disc, Wet Friction, Microstructure.

Accepted Manuscript

1. Introduction

Fuel consumption, crude oil depletion, emission legislative, friction losses and wear are considered vital parameters in the automotive industry constituting the true motivators. In the modern era there is extensive use of Diesel Engine in the field of transportation, power generation, agriculture and manufacturing. The life of an engine depends on the losses and wear experienced by the engine in terms of piston rings and cylinder liner. The piston rings and cylinder liner combine to form the tribopair that plays a significant role in power transformation from the chemical energy of the fuel in the combustion chamber. The largest source of frictional losses in the internal combustion engine is cylinder and piston system which comprises nearly 50% of the total frictional losses. The piston rings contribute 70 – 80% of friction in the cylinder and piston system. Thus, the cylinder and piston system being the paramount needs optimization to achieve high efficiency. In addition to frictional losses, this system also consumes a large volume of lubricating oil and releases a considerable amount of particulate emissions. In order to optimize the cylinder and piston system, an experimental investigation is of utmost importance to study its tribological behaviour.

Generally, the pin-on-disc method is considered to analyse the nature of materials in contact. The material of pin represents the cylinder liner while the disc material represents the piston rings. A comprehensive research has been done during the second half of the twentieth century for the development of more compact, fuel efficient and eco-friendly internal combustion engines for automobiles. Improvement in mechanical efficiency of internal combustion engines can be achieved by decreasing the frictional losses occurring in cylinder and piston assembly. Wear measurements had been carried out and analysed in cylinder and piston ring during their mating with each other when the automotive engine was operated under artificially created dusty environmental conditions [1 – 3]. Tests were conducted on pin-on-disc test apparatus with boundary-lubricated cast iron materials and the values of specific wear rates were observed to be between 10^{-13} and 10^{-10} $mm^3/mm/N$ [4]. Based on the Newton-Raphson-Murty algorithm, the non-linear finite-element method was used to analyse the piston rings problems based on the theories of elasto hydrodynamic and hydrodynamic lubrication [5]. The understanding of theoretical concepts and use of a methodology for estimating the service life of an I C engine, based on the law of adhesive wear and experimental evidence, were explained extensively [6]. The study and experimental research were conducted on materials, metallurgical aspects of piston ring scuffing, piston ring lubrication in a two-stroke diesel engine and the film thickness between the cylinder liner and piston rings [7 – 8]. Experiments were also performed on a reciprocating test rig having the supply of lubricant into the contact patch and measuring the friction force between specimen and counter-specimen. Tribological behaviors of cylinder liner with and without oil pockets were examined and compared [9 – 10]. A test method was developed for evaluation of the friction and wear of piston rings and cylinder liner used in heavy-duty diesel engines. These work focussed on the conditions of the lubricating oil and its effect [11 – 12]. At Volvo, technology had been updated for better evaluation of the friction between different combinations of materials. The elements under considerations were the piston rings and the cylinder liner which were analysed for friction, wear, and change in surface morphology during the running conditions [13]. The author investigated the role of sulfur for the high performance of the current tribo system. The field worn samples were examined to conclude that no tribo film containing sulphur existed under the tribological conditions set up for the piston rings and cylinder liner in the reciprocating test rig [14]. A similar work was carried out using the mathematical model to solve the Reynolds equation [15]. A non-firing engine simulator was developed which operated at same linear speed, load and stroke as that of an actual engine. Except firing pressures, all engine operating conditions were simulated [16]. Experimental investigation was carried out to study the characteristics of the cylinder liner-piston ring tribopair. The process parameters selected for the experiment were sliding velocity, applied load and type of lubricating oil. A mathematical model was developed using L16 Taguchi orthogonal design to optimize the reciprocating wear for minimum weight loss and friction. [17]. Author did an experimental study of single cylinder diesel engine operating at steady state to measure wear and surface roughness of cylinder liner by a WYKONT 1100 optical surface profile meter at various locations on the surface of cylinder liner before and after each test. It was observed that surface roughness decreased with time following which it experienced a steep decrease in rate during the run-in period. The results obtained in terms of volume of wear were reasonable and

Accepted Manuscript

within the tolerance limits [18]. An investigation of temperature effect on wear and friction was carried out when steel to steel contacts was un – lubricated, in various aerobic gaseous atmospheres and at high temperature of about 200°C. The inference was drawn considering the current results compared to previously published results [19]. A ring segment was tested in Laboratory to evaluate wear and friction behavior of piston ring on a flat specimen of grey cast iron representing cylinder liners in real engine. Various lubricants including mineral oil, Jet aviation fuel, fully formulated 15W40, category of heavy duty oil, were used to evaluate the condition of lubricant. During the test the temperatures were maintained from 25 to 100°C. An incremental load procedure was used to evaluate friction behavior which showed boundary lubrication for all lubricants at 100°C. Wear tests were performed at 240 N for 6 hours at a constant temperature of 100°C with new ring segments following which weight loss method was used to measure the wear. A geometric model was designed for wear volume and wear depth to take into account the compound curvatures before and after testing [23].

Carbon fabric was treated with gamma irradiation technique with varying doses (100–300 kGy) to develop composites with polyetherimide (PEI) matrix which was characterized for inter-laminar shear strength (ILSS) and adhesive wear studies. FTIR and SEM indicated presence of carbonyl functional group and roughening of surface inferred improved fiber–matrix interface [20]. Investigation was carried out to develop and analyse the material properties of a Polybenzimidazole (PBI) - Polyetherketone (PEK) blend based composite and composites containing solid lubricants (SLs), short fibers of glass and carbon. It was inferred from the experimental work that the SLs or fibers together showed excellent tribo- properties in accordance to the composite based on PBI–PEEK blend [21]. Experimentation was carried out for two diesel fuels i.e. EN590 and GDK650 as lubricants in the pin-on-disc model during the testing of wear behaviour of hydrogenated diamond like-carbon (DLC) coating in DLC/steel. It was concluded that anti-wear additives were harmful to DLC influencing its wear rate while steel remains unaffected by pressure, temperature, speed and type of lubricant used [22]. The investigation was conducted to analyse the tribological properties of ground, lapped and textured lapped discs against polished pins for reciprocating motions. Various output parameters were analysed among which coefficient of friction showed 30% reduction and vibrations between contact points of textured surface-balls interface showed a prominent dip at the resonance frequency in comparison to interface of polished surface-balls [24]. This work laid down the tribological analysis carried out for the dimpled piston rings. In order to curb the fuel consumption, research is being conducted to reduce the friction between cylinder liner and the piston rings. The experimental work yielded the result in favour of the dimpled piston rings as it helped saved considerable amount of fuel and lubricating medium. The tribological parameters, coefficient of friction and wear had been studied in this research work under full lubricating conditions for smooth/textured surfaces of steel discs and smooth surface of steel pins. Reduction in considered tribological parameters had been recorded in textured steel disc for fully flooded and starved conditions [25]. Another work investigated the mechanical properties of synthesised Mg_2Si co-doped with Bi and Sb through in situ spark plasma reaction sintering. The synthesis process displayed enhanced mechanical properties such as hardness ~5.4 ± 0.2 GPa and an elastic modulus ~142.5 ± 6 GPa with a fracture toughness of ~1.71 ± 0.1 MPa √m, t as thermal shock resistance as ~300 W m−1. It promised the sustainability of their thermoelectric modules under all circumstances [28]. Investigation was carried out on the stir casted aluminium based composite reinforced with 2–4–6% weight flyash. The pin-on-disc test rig was used to conduct the experiment for the smooth surfaces of cast iron disc and smooth MMC pin tribo-pair for the study of tribological parameters i.e. friction force and wear. The experimental work concluded that MMC with 6% flyash wears less while MMC with 4% flyash provides low coefficient of friction or friction force [28].

Based on the literature review, it is noticed that friction and wear reduction in piston and piston ring system are vital task. It has also been noticed that there is dearth of experimental studies towards exploring the tribo-dynamic behaviours of piston and piston ring system in situ condition under thin/mixed film lubrication. Hence, the objective of this paper is to present the experimental tribodynamic findings of fully flooded contacts of chrome steel pin

Accepted Manuscript

and cast iron / chrome plated cast iron discs in situ condition for highlighting the possibility of manufacturing of chrome coated liners and piston rings.

2. Experiments details

The specimens of the Cast Iron disc and Chromium-coated Cast Iron disc of diameter 165 mm and thickness 8mm were manufactured as depicted in the figure 1, for the conduction of experimental investigation. The surface finished discs attained the surface roughness level in the range 0.15 to 0.3 μm R.M.S. value. Surface roughness was measured by Telly Surf; 8 readings along the radial direction at interval of 45^{o} and 8 readings perpendicular to it from middle of the stroke of Telly Surf. A cylindrical Steel pin with diameter 8mm and length 70 mm was considered for testing the circular discs made of Cast Iron and Chromium-Coated Cast Iron. The face of the circular pin was leveled and checked by rubbing it against a soft paper for removing the debris if any. Face flatness was checked with the help of ink impression on paper. Pin was held rigidly 4 mm below the clamp for the test purpose on the disc to minimize bending and this distance was ignored while evaluating the tribological properties. Generally the lubricating oil (15W40) used in tropical regions was considered for the experimentation. The characteristics determined in the laboratory at 40^{O}C and atmospheric pressure were (1) kinematic viscosity of 160.32 mm^2/sec as per ASTM D – 445, (2) Density of 0.8744 g/cm^3 as per ASTM D – 4052 and (3) Specific gravity of 0.8752 as per ASTM D – 287 [26].

The schematic diagram of friction and wear tester used in the present study is shown in figure 2. It has a motor which can rotate at varying speeds based on the rheostat position. The desired speed was selected between 200 to 2000 r.p.m. There is a horizontal fixture on which the specimen disc was bolted at 4 points along with the speed sensor and the probe for measurement of wear. Weights for loading were placed on the pan which has a horizontal platform to support it in the direction of gravity. Weight placed in the pan is transmitted through a bell-crank mechanism comprising uni-lever via wire and pulleys. A force detector probe attached to the side of the bell crank to measure the shear friction force. Specifications of Friction and Wear tester used in these experiments were: Pin diameter 8 mm of chrome steel; Disc (i) diameter 165 mm and thickness 8 mm made of cast iron and disc (ii) diameter 165 mm and thickness 8 mm made of the chrome plated cast iron disc. According to the specifications of Tribometer the range of speed varies between 200 to 2000 r.p.m and frictional force can be measured upto 200 N with the least count of 0.1 N and Accuracy of $0.1 \pm 2\%$ measured frictional force in Newton (N).

In this experimental plan, the analysis has been designed in such a manner that each factor was analysed independently. Implementation of this plan showed significant reduction in the execution time needed for conducting the experiment. It was very efficient in investigating the effects of multiple performance factors and also to study the effect of individual factors. The experiment was performed in Tribological Lab, Delhi Technological University, Delhi. Pressure was varied between 1.95 MPa to 3.1 MPa and pin used for the same was 8 mm in diameter and 70 mm in length as shown in figure 1 and 2. Parameters were selected on the basis of preliminary investigations and 27 experiments were performed on each disc. Tribological study was carried out for the loads represented by pressure varying from 1.95 MPa to 1.4 MPa, at varying sliding speeds 2 – 10 m/s (equivalent to 200 – 2000 rpm) and distances traversed by the pin between 1000 m to 5000 m as per the experimental plan.

3. Results and discussions

*F*igure 3 shows a typical microstructural image of the cast iron disc and the corresponding elemental distribution maps, employing scanning electron microscope. Figure 4 shows the EDS spectrum, exhibiting the individual constituent elements in the cast iron disc, and the inset confirms the chemical composition of the individual elements. Figure 5 shows a microstructural image of the pin material and the corresponding elemental distribution maps, using scanning electron microscope. Figure 6 shows the EDS spectrum, depicting the individual constituent elements in the pin material, and the inset confirms the chemical composition of the individual elements. The morphology and composition were analyzed for the cast iron disc-steel pin tribopair which was then subjected to tribological performance employing Pin-on-Disc Tribometer (Make: Ducom, India). The results obtained have been discussed in the following paragraphs.

Accepted Manuscript

3.1 Coefficient of Friction

The trend of the coefficient of friction (μ) versus pressure (MPa) considering sliding speed of 2 m/s – 10 m/s has been plotted in figure 7 for the pressure ranging between 1.95 MPa – 3.1 MPa for flooded lubrication. Chromium plated disc shows improvement in the tribological parameters immensely in comparison to the cast iron disc. Experiments were performed by varying the distance between 1000 m to 5000 m at the same pressure range and in the presence of same lubricating oil (SAE 15W40). The coefficient of friction as in figure 7 was observed 0.03 to 0.04 between the steel pin and the cast-iron disc, that the plain cast iron disc had the maximum value of coefficient of friction i.e. 0.04. Under the same corresponding conditions, Chromium Plated cast iron disc had a coefficient of friction as 0.025 - 0.035. It was observed that coefficient of friction continuously decreased with an increase in pressure and attained the least value of coefficient of friction 0.025, at the pressure of 3.1 MPa. So the hard surface in relative motion had low coefficient of friction in fully flooded condition.

3.2 Wear

Wear characteristics were shown in figure 8, where initially fast growth in wear was observed at the start due to solid contact, interlocking of asperity and as tribopair had relative motion, breaking of the asperty took place as shown in figure 12. The variation of wear with different pressure values under the condition of fully flooded lubrication were displayed in Figure 8. A decrease in wear rate was observed as the pressure increased. This phenomenon was governed by thin layer mixed and elasto hydrodynamic lubrication. Under contact stresses, a very thin film of lubricating oil formed that separate the contact surfaces. It resulted in elasto hydrodynamic lubrication. This resulted in reduction of friction and wear. Wear pattern on the pin material was shown in figure 12 (a). The initial increase in wear occurred due to the abrasion wear owing to the highly strained fragments in between the contact surfaces and later decrease was due to the diffusion of surface debris which was insignificant in comparison to the abrasion wear.

The decrease in wear was also due to the formation of thin layer of lubricant at the interface of the discs and pin. This resulted into thin layer mixed lubrication processes. In experiment, the various values of distances were taken from 1000 - 5000 m for the disc with predefined diameter against varying pressure values i.e. 1.95 MPa, 2.5 MPa, and 3.1 MPa. Under fully flooded lubrication, the wear increased from 1×10^{-10} m^3 to 15×10^{-10} m^3, was due to thin layer mixed and elasto hydrodynamic lubrication. Hence, it was inferred that the pin material wears more as the pressure is increased.

Microscopic views as in figure 12 and extended view as in figure 1 depict the abrasive wear and ploughing took place in thin layer lubrication

3.3 Vibrations Amplitude

Experiments were conducted taking various values of distances from 1000 m to 5000 m at different disc diameters and different pressure. It was observed that when lubrication was fully flooded, the vibration amplitude varied between 40 μm and 52 μm as shown in figure 9. It was due to the hydrodynamic lubrication. As per figure 9, the amplitude of vibrations increases with the pressure at the considered values i.e.1.95 MPa, 2.5 MPa, and 3.1 MPa. This was caused by excessive pressure at the interface of the tribopair. The vibrational values showed the experiment carried out was stable.

3.4 Temperature Rise

Continuous relative motion between two metals is the source of temperature and as the motion continues the temperature rises continuously. The rise in temperature was governed by the surface roughness of the materials in contact along with the mechanical properties of the materials of pin and disc. Experiments were conducted at varying distance travelled in the range of 1000 -

Accepted Manuscript

5000 m for different diameter of the disc with the pressure values as 1.95 MPa, 2.5 MPa and 3.1 MPa. Figure 10 depicts the temperature rise at the interface of the tribopair and it varies between 4 K and 9 K as measured by thermal image camera when lubricant was flowing continuously. Observed temp rise value showed the experiment was conducted with the stable pace and energy loss was very less

3.5 Sound

The rules of noise pollution, 2000, had set different acceptable limits depending upon the category i.e. day and night. For industries, it is 75 dB for the day and 70 dB for the night. Similarly, for residential areas, it is 55 dB for the day and 45 dB for the night. Figure 11 shows the trend of sound as observed during the experimentation procedure. It was noted in the Figure 11 that sound varies from 66.8 dB to 78 dB. High noise observed in the Chromium plated disc as compared to the Cast Iron disc. This was due to the property of cast iron of absorbing the energy or vibration experienced during the running conditions. At low speeds and for longer distance traversed it was ascertained that cast iron disc absorbed more energy and hence made less noise. The continuous flow of lubricant also acted as the medium of sound for the base material in motion with respect to pin material. The measuring instrument was the sound meter which was kept 0.25 m away from the pin and disc.

4. Conclusion

1. It has been observed that Chromium plated cast iron disc significantly improved the coefficient of friction and the value 0.025 found to be at 3.1 MPa and 6m/s relative velocity at the interface under the condition of fully flooded lubrication as compared with cast iron disc with chrome steel pin. In such condition, the lubricating oil helped in reducing the shear force which further reduced the coefficient of friction with the increase in speed. The chrome plated cast iron disc had hard surface to reduce coefficient of friction at thin layer tribo lubrication.
2. Chrome plated cast iron disc experienced less wearing due to the formation of hard surface layer of chromium and the presence of thin layer mixed and elasto hydrodynamic lubrication in the contact area of the tribopair. The chrome plated cast iron disc had hard surface to reduce wear rate at thin layer tribo lubrication.
3. The amplitude of vibration was observed to be minimum at 40 μm and maximum at 52 μm. The Chromium Plated cast iron disc had more amplitude of vibration in comparison to cast iron disc due to less energy absorption in the chromium plated cast iron disc.
4. The rise in temperature at the interface of the tribo - pair formed between the cast Iron, Chromium plated cast Iron discs and the steel pin was observed from 4 K to 9 K. The combination of chromium plated cast iron disc with steel pin showed more temperature rise when compared with cast iron disc and steel pin. This rise in temperature was negligible due to the condition of fully flooded lubrication.
5. The experimental results show that sound varied between 66.8 dB to 78 dB. Chromium plated cast iron disc owing to hard surface generated more noise while the cast iron disc generated less noise due to its property of absorption of energy and also the continuous flow of lubrication acted as the sound absorber leading to a quiet movement between the two surfaces in contact.

Acknowledgement

Authors would like to thank Nanoscale Research Facility (NRF) of IIT Delhi, India and Tribology Lab of Delhi Technological University, Delhi, India for their kind help in characterization.

Accepted Manuscript

References

[1] Wakuri Y, Hamatake T, Soejima M, Kitahara T 1992 Piston Ring Friction in Internal Combustion Engines *Tribol. Int*. 25 299-308

[2] Nautiyal P C, Singhal S, Sharma J P 1983 Friction and wear processes in piston rings *Trobol. Int*. 16 43-49

[3] Pawlus P 1993 Effects of honed cylinder surface topography on the wear of piston-piston ring-cylinder assemblies under artificially increased dustiness conditions *Tribol. Int*. 26 49-55

[4] Childs T H C and Sabbagh F 1989 Boundary-lubricated wear of cast irons to simulate automotive piston ring wear rates *Wear,* 134 81 – 97.

[5] Hwu C J and Weng C 1991 Elastohydrodynamic lubrication of piston rings *Wear* 150 203-15

[6] Picken D J, Hassaan H A 1983 A method for estimating overhaul life of internal combustion engines including engines operating on biogas and methane *J. Agric. Engg. Res.* 28 139-47

[7] Scott D, Smith A I, Tait J and Tremain G R 1975 Materials and metallurgical aspects of piston ring scuffing - A literature survey *Wear.*33 293–315

[8] Moore S L 1981 Piston ring lubrication in a two-stroke diesel engine *Wear* 72 353-69

[9] Grabon W, Pawlus P, Sep J 2010 Tribological characteristics of one-process and two-process cylinder liner honed surfaces under reciprocating sliding conditions *Tribol. Int.* 43 1882–92

[10] Grabon W, Koszela W, Pawlus P, Ochwat S 2013 Improving tribological behaviour of piston ring–cylinder liner frictional pair by liner surface texturing *Tribol. Int.* 61 102–08

[11] Truhan J J, Qu J, Blau P J 2005 The effect of lubricating oil condition on the friction and wear of piston ring and cylinder liner materials in a reciprocating bench test *Wear* 259 1048–55

[12] Michalski J, Wo P 2011 The effect of cylinder liner surface topography on abrasive wear of piston–cylinder assembly in combustion engine *Wear* 271 582–89

[13] Johansson S, Nilsson P H, Ohlsson R, Rosén B G 2011 Experimental friction evaluation of cylinder liner/piston ring contact *Wear* 271 625–33

[14] Olander P, Hollman P, Jacobson S 2013 Piston ring and cylinder liner wear aggravation caused by transition to greener ship transports–Comparison of samples from test rig and field *Wear* 302 1345–50

[15] Mezghani S., Demircia I., Zahouani H., Mansori M. El 2012 The effect of groove texture patterns on piston-ring pack friction *Precision Engineering* 36 210– 217

[16] Srivastava D K, Agarwal A K, Kumar J 2007 Effect of liner surface properties on wear and friction in a non-firing engine simulator *Materials and Design* 28 1632–40

[17] Kapsiz M, Durat M, Ferit F 2011 Friction and wear studies between cylinder liner and piston ring pair using Taguchi design method *Advances in Engineering Software* 42 595–603

[18] Gara L, Zou Q, Sangeorzan B P, Barber G C, McCormick H E, Mekari M H 2010 Wear measurement of the cylinder liner of a single cylinder diesel engine using a replication method", *Wear* 268 558–564

[19] Velkavrh I, Ausserer F, Klien S, Voyer J, Ristow A, Brenner J, Forêt P, Diem A 2016 The influence of temperature on friction and wear of un-lubricated steel/steel contacts in different gaseous atmospheres *Tribol. Int.* 98 155–171

[20] Tiwari S, Bijwe J, Panier S 2011 Gamma radiation treatment of carbon fabric to improve the fiber–matrix adhesion and tribo-performance of composites *Wear* 271 2184–2192

[21] Bijwe J, Gupta M K, Parida T, Trivedi P 2015 Design and development of advanced polymer composites as high performance tribo-materials based on blends of PEK and ABPBI *Wear* 342-343 65–76

[22] Djoufack M H, May U, Repphun G, Brögelmann T, Bobzin K 2015 Wear behaviour of hydrogenated DLC in a pin-on-disc model test under lubrication with different diesel fuel types *Tribol. Int*. 92 12–20

Accepted Manuscript

[23] Truhan J J, Qu J and Blau P J 2005 A rig test to measure friction and wear of heavy duty diesel engine piston rings and cylinder liners using realistic lubricants *Tribol. Int.* 38 211–218
[24] Sudeep U, Pandey R K and Tandon N 2013 Effects of surface texturing on friction and vibration behaviours of sliding lubricated concentrated point contacts under linear reciprocating motion *Tribol. Int.* 62 198–207
[25] Pandey R K, Tandon N, Singh A K 2013 Fuel saving in IC engine by surface texturing of piston rings *Nat. Conf. on Recent Advancements in Mech. Engg. (NCRAME-2013) 8-9 November 2013 NERIST, Itanagar, Arunachal Pradesh, India* 1-4
[26] Singh R C, Pandey R K, Rooplal, Ranganath M S and Maji S 2016 Tribological performance analysis of textured steel surfaces under lubricating conditions *Surf. Topogr.: Metrol. Prop*. 4 1-9
[27] Sharma V K, Singh R C, Chaudhary R 2017 Effect of fly ash particles with aluminium melt on the wear of aluminium metal matrix composites *Engg. Sc. and Tech., an Int. J.* Article in press 1-6
[28] Muthiah S, Singh R C, PathakB D and Dhar A 2017 Mechanical properties of thermoelectric n-type magnesium silicide synthesized employing *in situ* spark plasma reaction sintering *Mater. Res. Express* 4 1-8

Accepted Manuscript

Figure captions:

Figure 1. Schematic and Dimensional Details of Disc and Pin

Figure 2. Schematic of Friction and Wear Tester

Figure 3. (a) Scanning electron images of (a) Cas iron disc and elemental mappings of (b) Iron, (c) Carbon and (d) Silicon

Figure 4. EDS spectrum showing the elements of Fe, C and Si presence in the disc (Inset shows the chemical stoichiometric of the individual elements)

Figure 5. (a) Scanning electron images of (a) Pin material and elemental mappings of (b) Oxygen, (c) Iron, (d) Silicon, (e) Carbon, and (f) Chromium

Figure 6. EDS spectrum showing the elements of Fe, C, O, Cr and Si presence in the pin (Inset shows the chemical stoichiometric of the individual elements)

Figure 7. Coefficient of friction (μ) of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Figure 8. Wear loss of Steel pin against Cast Iron/ Chromium plated Cast Iron with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Figure 9. Vibration amplitude of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Figure 10. Temperature rise of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Figure 11. Frictional sound of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Figure 11.Microscopic view of worn surface of (a) and (b) cast iron disc.

Accepted Manuscript

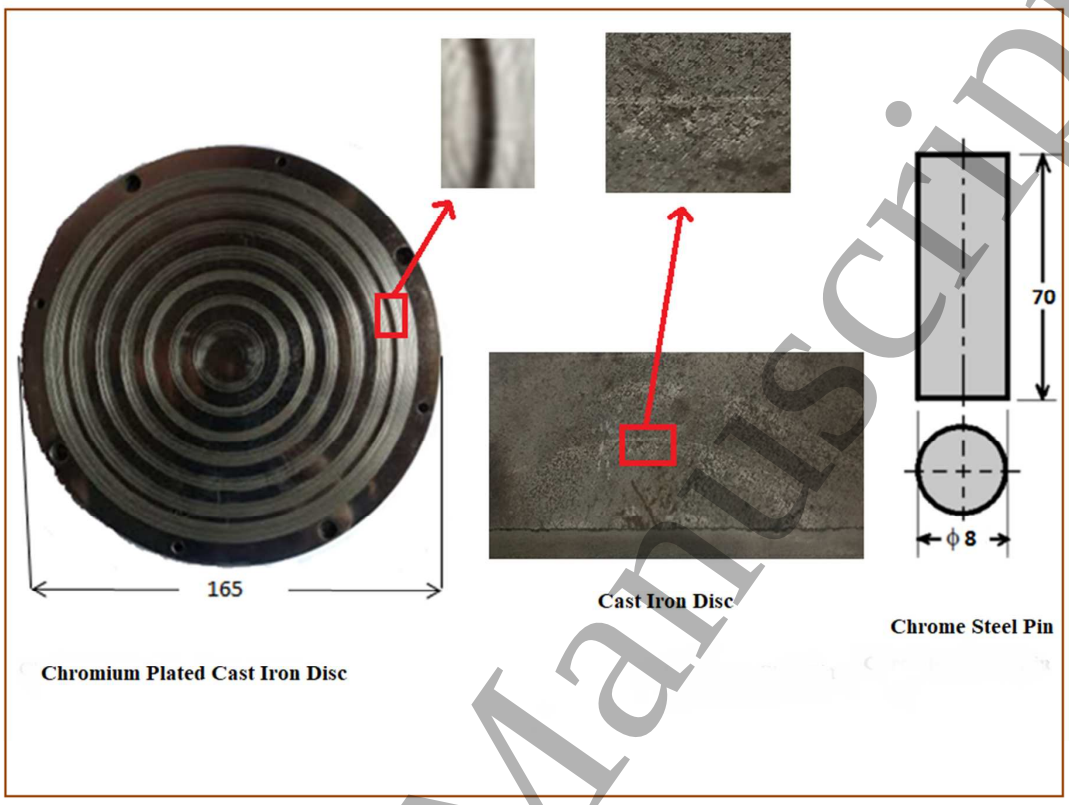


Figure 1. Schematic and Dimensional Details of Disc and Pin

Accepted Manuscript

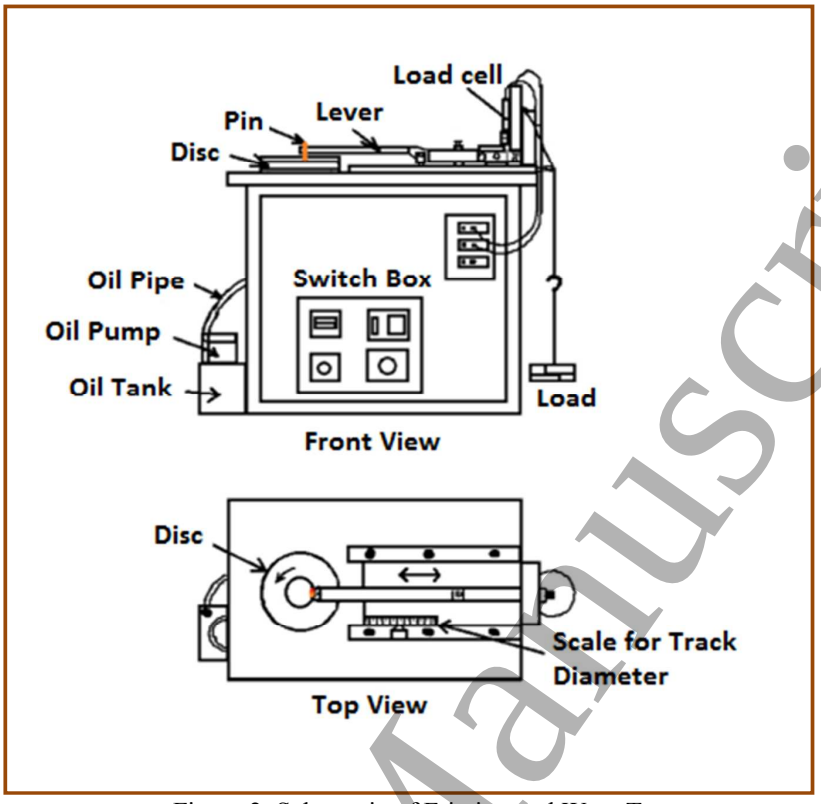


Figure 2. Schematic of Friction and Wear Tester

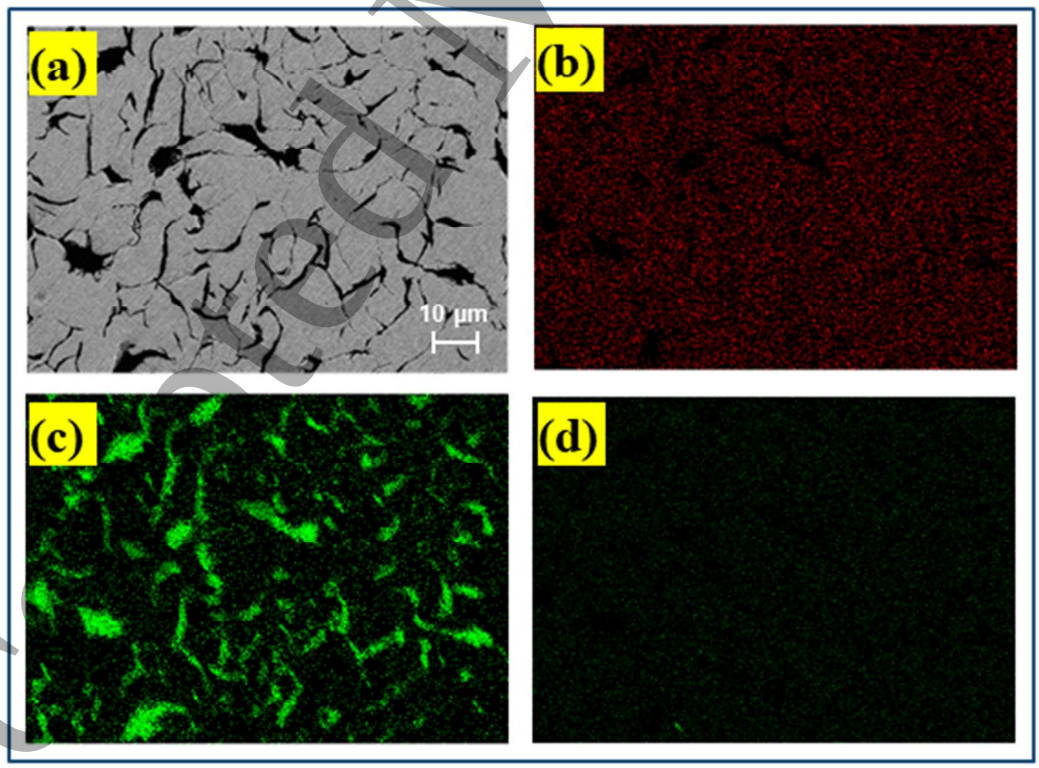


Figure 3. (a) Scanning electron images of (a) Cas iron disc and elemental mappings of (b) Iron, (c) Carbon and (d) Silicon

Accepted Manuscript

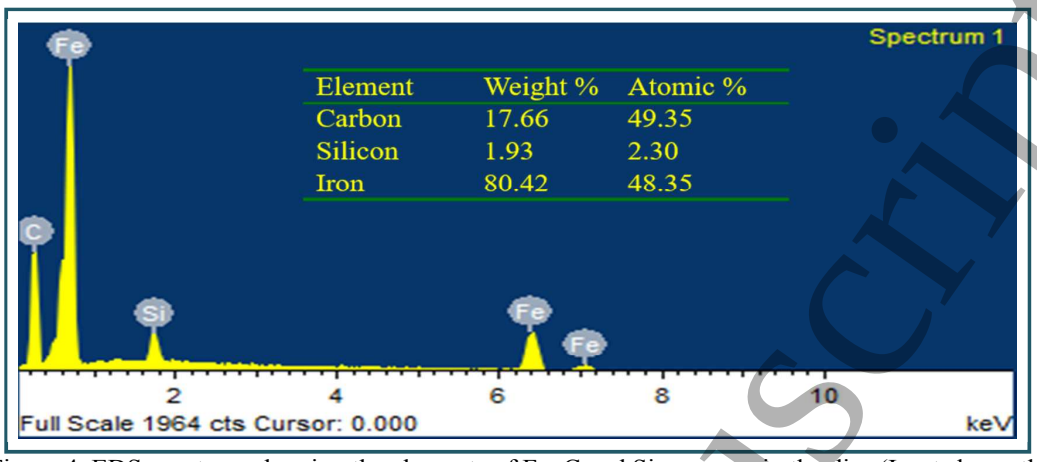


Figure 4. EDS spectrum showing the elements of Fe, C and Si presence in the disc (Inset shows the chemical stoichiometric of the individual elements)

Accepted Manuscript

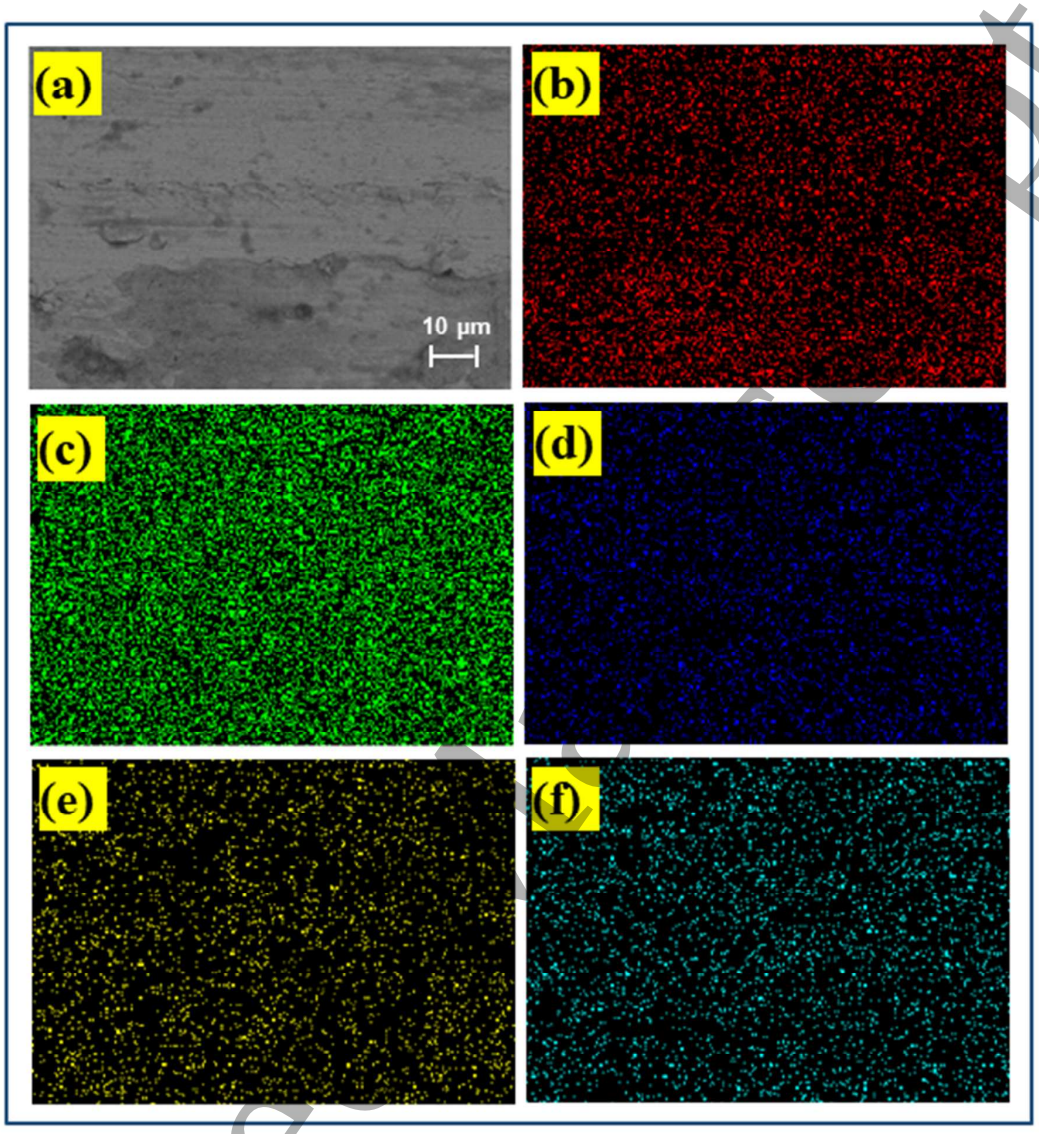


Figure 5. (a) Scanning electron images of (a) Pin material and elemental mappings of (b) Oxygen, (c) Iron, (d) Silicon, (e) Carbon, and (f) Chromium

Accepted Manuscript

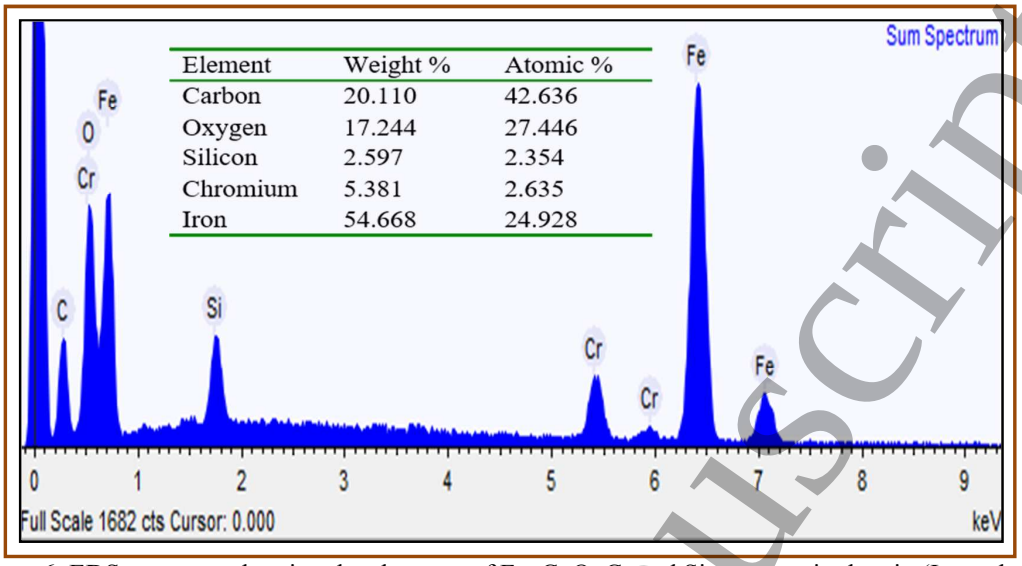


Figure 6. EDS spectrum showing the elements of Fe, C, O, Cr and Si presence in the pin (Inset shows the chemical stoichiometric of the individual elements)

Accepted Manuscript

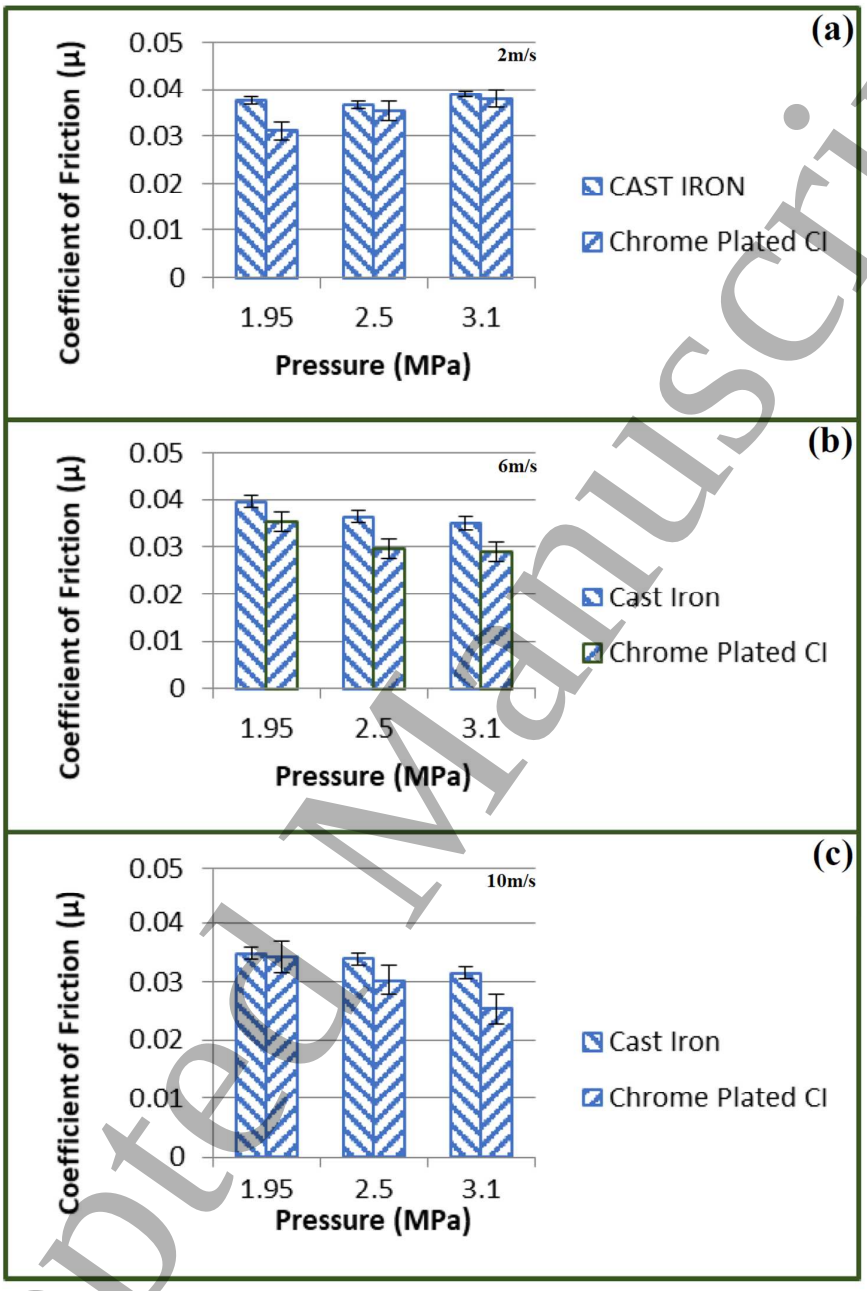


Figure 7. Coefficient of friction (μ) of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Accepted Manuscript

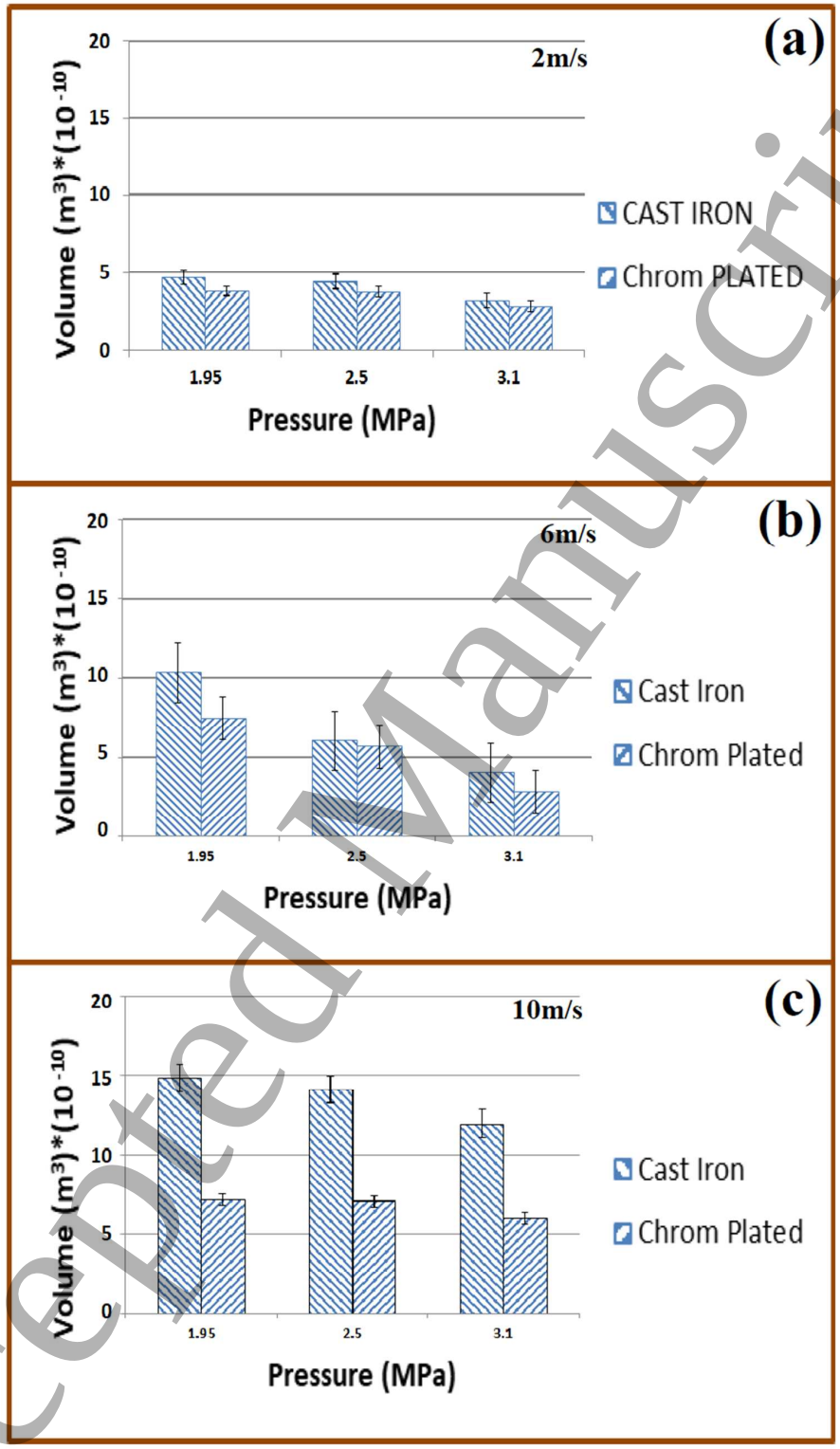


Figure 8. Wear loss of Steel pin against Cast Iron/ Chromium plated Cast Iron with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Accepted Manuscript

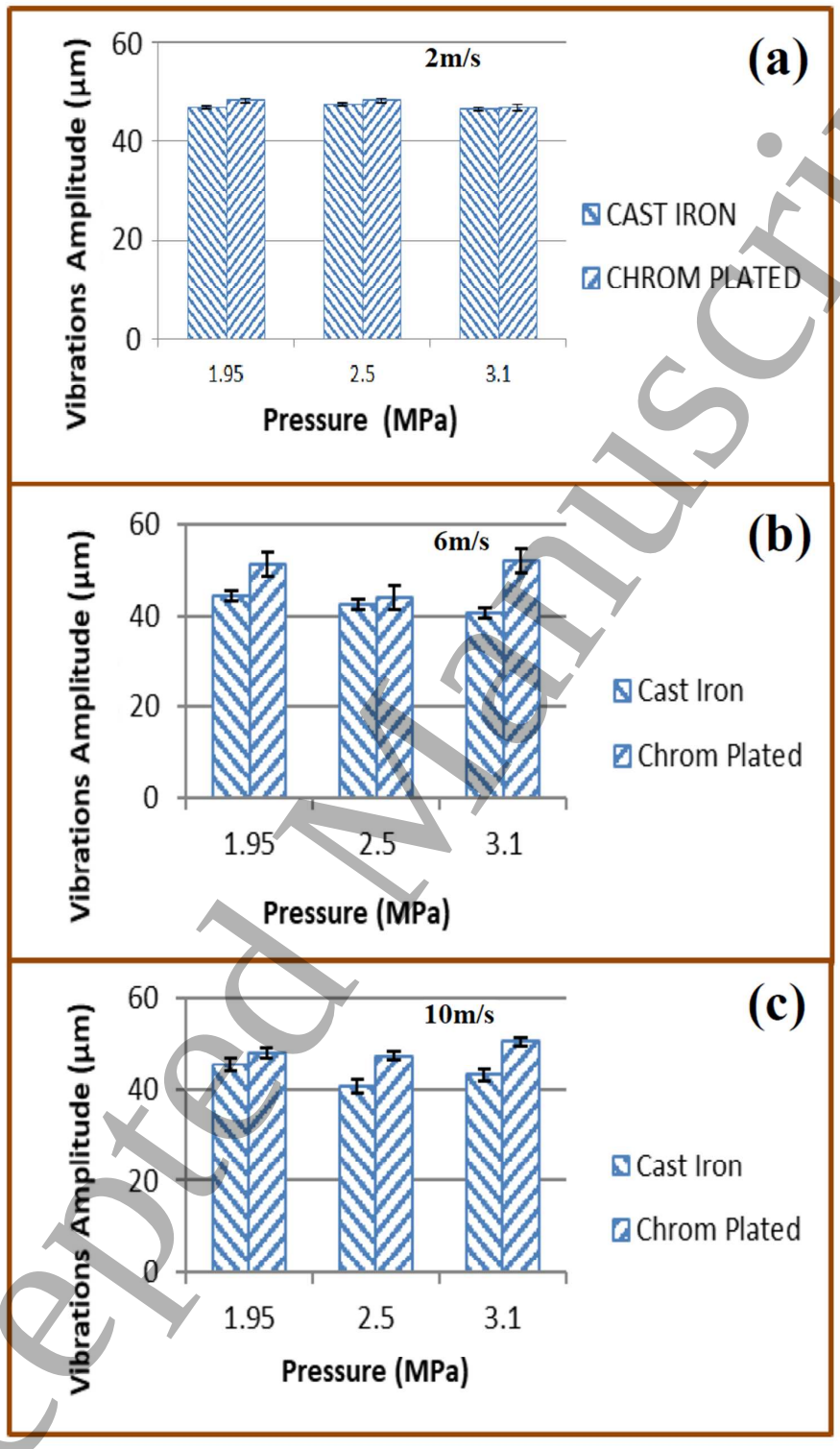


Figure 9. Vibration amplitude of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Accepted Manuscript

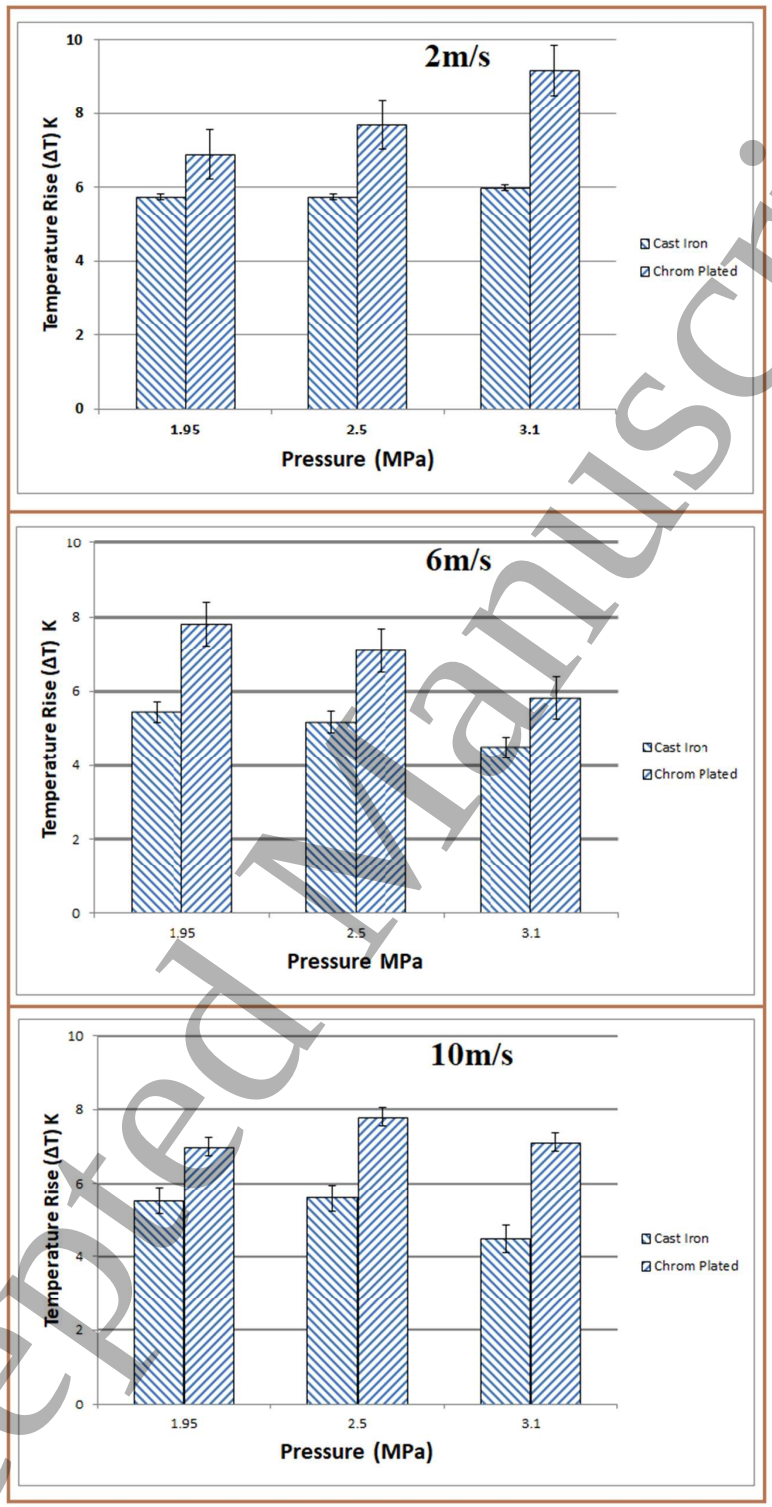


Figure 10. Temperature rise of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Accepted Manuscript

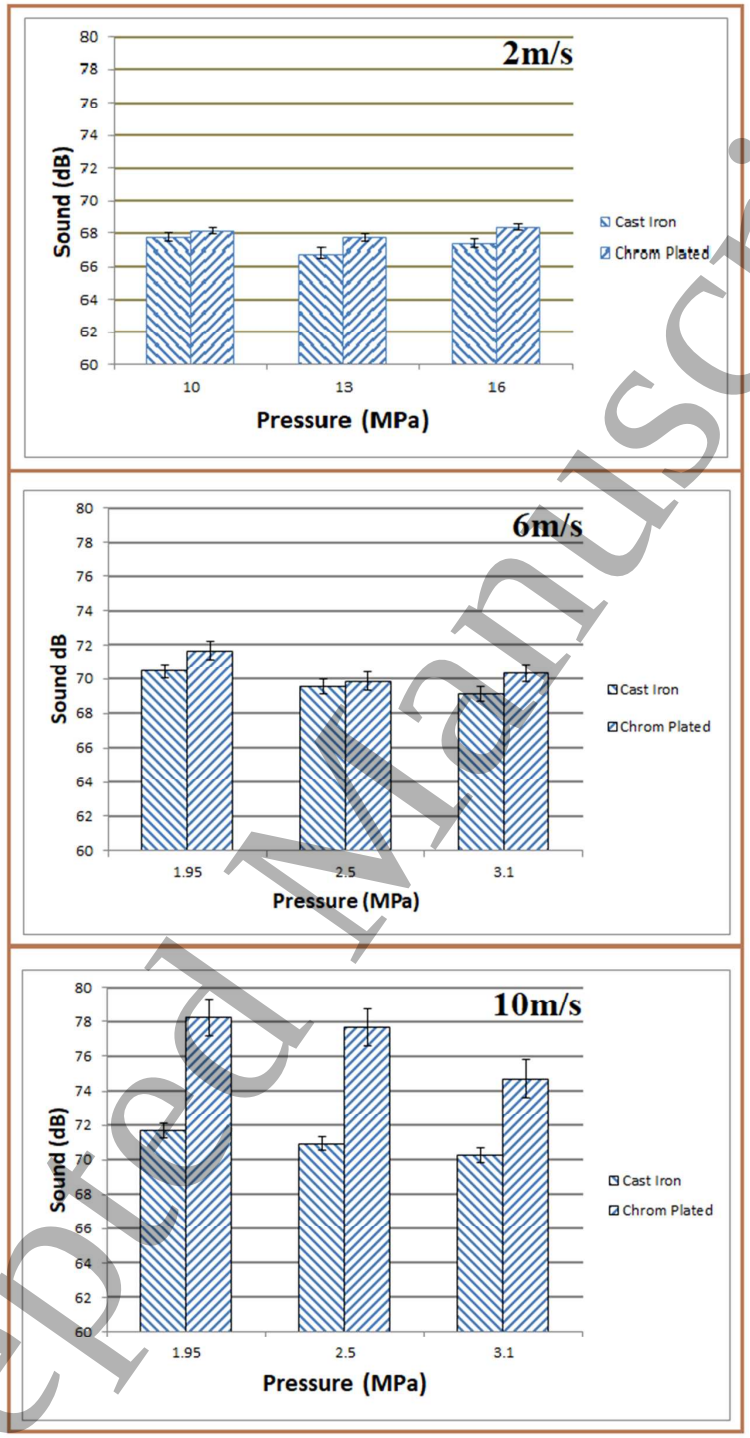


Figure 11. Frictional sound of Cast Iron/ Chromium plated Cast Iron against Steel pin with three different pressure (1.95, 2.5, 3.1 MPa) and speed (a) 2 m/s, (b) 6 m/s and, (c) 10 m/s conditions

Accepted Manuscript

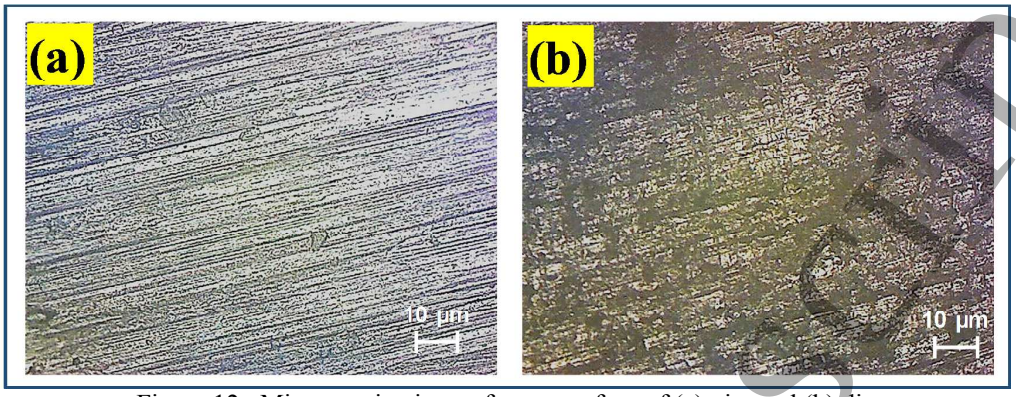


Figure 12. Microscopic views of worn surface of (a) pin, and (b) disc

Accepted Manuscript

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325263090

Experimental Validation of PID and LQR Control Techniques for Stabilization of Cart Inverted Pendulum System

Conference Paper · May 2018

CITATIONS
0

READS
6

4 authors, including:



Sudarshan K Valluru
Delhi Technological University (Formerly Delhi College of Enginee...
23 PUBLICATIONS **21** CITATIONS
SEE PROFILE



Mayank Singh
Delhi Technological University
1 PUBLICATION **0** CITATIONS
SEE PROFILE



Vanshaj Khattar
Delhi Technological University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

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

meta algorithms application to nonlinear dynamical systems View project

meta algorithms applications to nonlinear dynamical systems View project

All content following this page was uploaded by Sudarshan K Valluru on 21 May 2018.

The user has requested enhancement of the downloaded file.

Experimental Validation of PID and LQR Control Techniques for Stabilization of Cart Inverted Pendulum System

Sudarshan K.Valluru
Incubation Centre for Control,
Dynamical systems, and Computation
Electrical Engineering Department
Delhi Technological University
Delhi-110042, India
sudarshan_valluru@dce.ac.in

Madhusudan Singh
Incubation Centre for Control,
Dynamical systems, and Computation
Electrical Engineering Department
Delhi Technological University
Delhi-110042, India
madhusudan@dce.ac.in

Mayank Singh
Incubation Centre for Control,
Dynamical systems, and Computation
Electrical Engineering Department
Delhi Technological University
Delhi-110042, India
mayank202244@gmail.com

Vanshaj Khattar
Incubation Centre for Control,
Dynamical systems, and Computation
Electrical Engineering Department
Delhi Technological University
Delhi-110042, India
vanshaj811@gmail.com

***Abstract*— This paper describes the design and implementation of control strategies for the stabilization of a Cart Inverted Pendulum system. An experimental model of the system is considered for the design of controllers. The real-time implementation is done by employing a two-loop PID controllers for the system where the gain values are pre-determined by using Frequency Response and Linear-Quadratic Regulator(LQR) methods. A comparison is made between the proposed design strategies in terms of steady state and transient performance parameters. It is observed that Linear-Quadratic method exhibits superior performance as compared to frequency response method.**

Keywords*— Cart Inverted Pendulum system; PID Controller; LQR Controller; Full State Feedback

I. INTRODUCTION

The Inverted Pendulum System is among the most complex systems for its inherent instability and is one of the benchmark control problems given by international federation of automatic control theory committee (IFAC)[1]. The control of this system has been a thoroughly researched problem in the field of systems engineering and is now considered as a benchmark system to develop new control strategies[2]. The inverted pendulum belongs to a class of underactuated mechanical system which cannot be commanded to follow arbitrary trajectories. It has fewer inputs than the total degree-of-freedom (DOF), and it has two DOF, one for cart motion in the horizontal plane, and another for the angular motion of the pendulum in the vertical plane. However, only the cart position is actuated, and the pendulum is indirectly affected/controlled. This experimental study aims to stabilize the cart inverted pendulum system and pendulum erected in an upward direction such that the position of the cart on the track is controlled in short interval time. The setup of the experimental system[3] is shown in Fig.1, which can be used to determine control laws or control strategies to achieve the desired system response and performance. The performance of the dynamical systems being controlled is desired to be an optimal response by many optimal control strategies such as PID, LQR [4], etc.

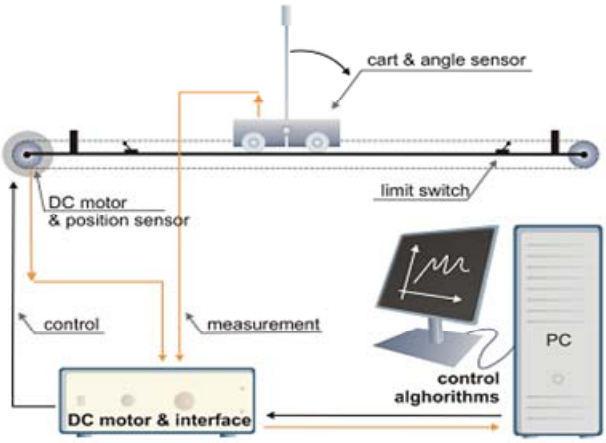


Fig. 1 Experimental setup of cart inverted pendulum system [3]

This paper uses the performance index criteria under the input voltage constraint, and implement the real-time experiment using the obtained optimal conditions. The real-time implementation is performed with two-loop PID controller, by using frequency response analysis and then using LQR optimal control. The two-loop PID controller[5] is designed by deriving the combined closed loop characteristic equation, and then the PID values are obtained by comparing it with the stable closed loop characteristic equation. This stable closed loop characteristic equation is obtained by adding a compensator and changing its parameters until to obtain the desired time and frequency response characteristics with a gain

978-1-5386-2440-1/18/$31.00 ©2018 IEEE

margin of greater than 2 and phase margin are higher than 60 degrees[6].

This paper organized in six sections commencing from Introduction followed by Section II which discusses the mathematical modeling of the system. Section III discusses the PID controller design using frequency response method. Section IV presents the Linear Quadratic Regulator design method using optimal control method. Section V includes the results of both methods and comparison is presented and ends with a conclusion as Section VI.

II. Modeling of Cart Inverted Pendulum System

Mathematical equations are derived using Newtonian mechanics. The system free body diagram is represented in Fig.2, and the corresponding non-linear equations are derived. The inverted pendulum experimental set-up system parameters are given in Table.I.The set-up is computer controlled and any desired controller to stabilize the system can be implemented through SIMULINK in MATLAB. The feedback signals to the computer and control signals from the computer are all given through the PCI- 1711 A/D card manufactured by Advantech Technologies.

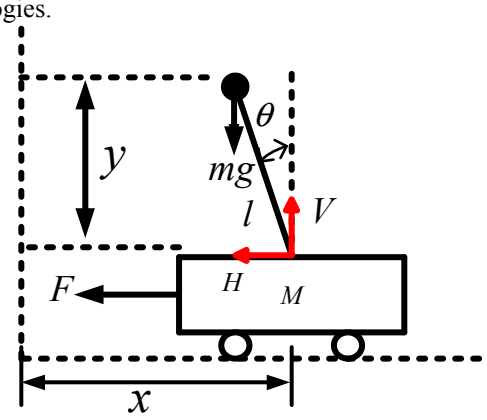


Fig. 2 Free body diagram of Cart Inverted Pendulum System

TABLE I. Cart Inverted Pendulum System Parameters

Parameter	Value
F - Force applied to cart	±24N
x - Cart position from the reference	±0.5 m
θ - Pendulum angle w.r.t verticle axis	0.1 rad
g - Gravity	9.81m/s^2
l - Length of the pole	0.38m
M - Cart mass	2.4 Kg
m - Pole mass	0.23 Kg
I - Moment of inertia of pole	0.099Kg/m^2
b - Cart friction coefficient	0.05Ns/m
d - Pendulum damping coefficient	0.05 s /rad

Newtonian mechanics is used to developing the nonlinear equations for the system. Forces along translational motion and rotational motion are taken into consideration, and thus two coordinate systems are taken into account from Fig.2.

H -Net horizontal force on the cart

V - Net vertical force on the pendulum

θ - Pendulum angle w.r.t vertical axis

mg - Weight of the bob

τ - Torque about the center of gravity of the pendulum

$$m\frac{d^2}{dt^2}\left(x - l\sin\theta\right) = -H \tag{1}$$

$$m\frac{d^2}{dt^2}\left(l\cos\theta\right) = V - mg \tag{2}$$

Taking moments about the center of gravity of the pendulum,

$$\tau_{net} = I\frac{d^2\theta}{dt^2} \tag{3}$$

$$-H\cos\theta + V\sin\theta - d\dot{\theta} = I\ddot{\theta} \tag{4}$$

From the equations (1) and (2), the pole equation is derived as (5)

$$\left(-ml\cos\theta\right)\ddot{x} + \left(I + ml^2\right)\ddot{\theta} + d\dot{\theta} - mgl\sin\theta = 0 \tag{5}$$

The cart dynamic equations are expressed as

$$F_{net} = M\ \ddot{x} \tag{6}$$

$$F + H - b\dot{x} = M\ \ddot{x} \tag{7}$$

$$(M+m)\ddot{x} + b\dot{x} - (ml\cos\theta)\ddot{\theta} + (ml\sin\theta)(\dot{\theta})^2 = F \tag{8}$$

Where equations (5) and (8) represent the nonlinear equations of the system and can be written as equation (9),

$$\begin{bmatrix}\ddot{x}\\ \ddot{\theta}\end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix}-ml\cos\theta & -(I+ml^2)\\ -(M+m) & -ml\cos\theta\end{bmatrix}\begin{bmatrix}mgl\sin\theta - d\dot{\theta}\\ F - \left(ml\sin\theta\right)\left(\dot{\theta}\right)^2 - b\dot{x}\end{bmatrix} \tag{9}$$

Where $\Delta = m^2l^2\cos\theta - \left(M+m\right)\left(I+ml^2\right)$

$$\begin{bmatrix}\ddot{x}\\ \ddot{\theta}\end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix}ml\cos\theta\left(d\dot{\theta} - mgl\sin\theta\right) + \left(I+ml^2\right)(ml\sin\theta(\dot{\theta}^2) + b\dot{x}\\ \left(M+m\right)\left(d\dot{\theta} - mgl\sin\theta\right) + \left(ml\cos\theta\right)(ml\sin\theta(\dot{\theta}^2) + b\dot{x}\end{bmatrix} + \frac{1}{\Delta}F\begin{bmatrix}-\left(I+ml^2\right)\\ -\left(ml\cos\theta\right)\end{bmatrix} \tag{10}$$

The state vector taken as $X = \left[x_1, x_2, x_3, x_4\right]^T = \left[x, \dot{x}, \theta, \dot{\theta}\right]^T$ and linearizing the non-linear differential equation for the cart inverted pendulum system at the unstable equilibrium point $\theta = 0$, and by taking $\sin\theta \approx \theta$, $\cos\theta \approx 1$, and $\dot{\theta}^2 \approx 0$. Thus, the linearized state-space model can be written as,

$$\dot{X} = \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & \frac{\left(I+ml^2\right)b}{\Delta} & \frac{-m^2gl^2}{\Delta} & \frac{mld}{\Delta}\\ 0 & 0 & 0 & 1\\ 0 & \frac{mlb}{\Delta} & \frac{-mgl\left(M+m\right)}{\Delta} & \frac{d\left(M+m\right)}{\Delta}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix} + F\begin{bmatrix}0\\ -\frac{(I+ml^2)}{\Delta}\\ 0\\ -\frac{ml}{\Delta}\end{bmatrix} \tag{11}$$

Where Δ is $(ml)^2 - \left(I+ml^2\right)(M+m)$. The state-variables of the continuous-time linear dynamical system are chosen

as $x, \dot{x}, \theta, \dot{\theta}$ where x is the position of the cart and θ is the angle of the pendulum, which are described by equations (12) and (13).

$$\dot{x} = Ax + Bu \tag{12}$$

$$y = Cx \tag{13}$$

By substituting the cart inverted pendulum system parameters, the obtained A, B, C and D values are given in (14), (15),(16) and (17).

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -0.02 & 0.22 & -0.0013 \\ 0 & 0 & 0 & 1 \\ 0 & -0.013 & 6.632 & -0.034 \end{bmatrix} \tag{14}$$

$$B = \begin{bmatrix} 0 \\ 0.388 \\ 0 \\ 0.257 \end{bmatrix} \tag{15}$$

$$C = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{16}$$

$$D = \begin{bmatrix} 0 \end{bmatrix} \tag{17}$$

The transfer functions of the cart and pendulum are shown in equation (18) and (19)

$$T_x = \frac{X}{F} = \frac{0.132\left(s^2 + 0.0378s - 12.98\right)}{0.339\left(s^3 + 0.058s^2 - 13.27s - 0.253\right)} \tag{18}$$

$$T_\theta = \frac{\theta}{F} = \frac{0.0874s^2}{0.339\left(s^3 + 0.058s^2 - 13.27s - 0.253\right)} \tag{19}$$

III. Design of PID controllers for cart inverted pendulum system

The designed two loop control strategy structure diagram for cart inverted pendulum system is shown in Fig.3.

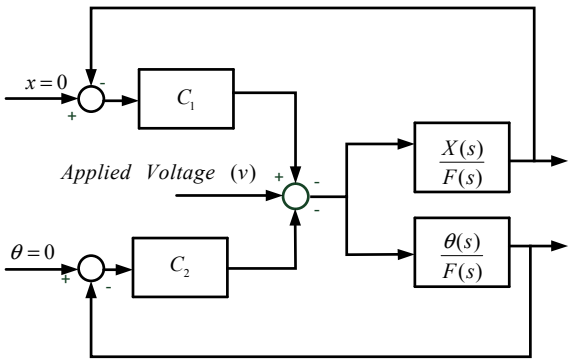


Fig. 3 Closed-loop PID structure diagram

$$C_1 = K_{p1} + \frac{K_{i1}}{s} + K_{d1}s \tag{20}$$

$$C_2 = K_{p2} + \frac{K_{i2}}{s} + K_{d2}s \tag{21}$$

Moreover, the closed loop transfer functions are obtained from Fig.3.

$$\begin{bmatrix} \frac{X}{F} \\ \frac{\theta}{F} \end{bmatrix} = \frac{\begin{bmatrix} T_1 \\ T_2 \end{bmatrix}}{1 + T_1C_1 + T_2C_2} \tag{22}$$

Where F is the applied force, T_1 and T_2 are the respective transfer functions of x and θ respectively, $1 + T_1C_1 + T_2C_2$ is the characteristic equation of the closed loop system. After substituting the equations (20) and (21) in (22), the net characteristic equation for the closed-loop is rewritten in equation (23).

$$s^5 + \left(0.388K_{d1} + 0.257K_{d2} + 0.0581\right)s^4 + \left(0.0147K_{d1} + 0.388K_{p1} + 0.257K_{p2} - 6.63\right)s^3 + \left(0.388K_{i1} + 0.257K_{i2} - 2.52K_{d1} + 0.0147K_{p1} - 0.126\right)s^2 + \left(0.0147K_{i1} - 2.52K_{p1}\right)s - 2.52K_{i1} = 0 \tag{23}$$

The Nyquist plot of the pendulum is shown in Fig. 4, which is used to tune the PID controller, where a compensator is assumed to be in parallel, the poles and zeros are added to the compensator until we get the desired frequency and time response. Nyquist stability criterion, that is Z=P+N, where Z gives the closed loop poles in the right half plane, P is the number of zeros of open loop transfer function, and N is the number of encirclements of -1 of the Nyquist plot in the clockwise direction. Z should be zero for the closed loop system to be stable.

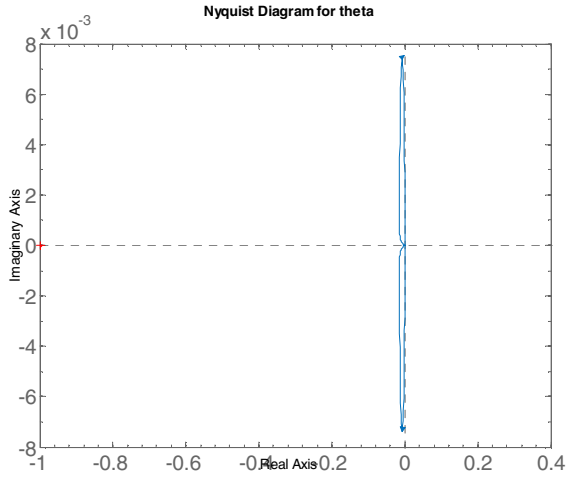


Fig.4 Nyquist plot of pendulum angle

The desired compensator transfer function is obtained as

$$T_{ctr} = \frac{200s^2 + 920s + 133.5}{s} \tag{24}$$

The net transfer function of the closed loop is calculated from the above compensator which gives a stable system. The obtained characteristic equation can be compared with equation (23) to obtain the PID values for both the controllers. The net closed loop characteristic equation of the closed loop system is assumed as

$$s^5 + p_1s^4 + p_2s^3 + p_3s^2 + p_4s + p_5 = 0 \tag{25}$$

To find the least square norm solution, let coefficient matrix be M, controller parameter matrix be P, and the stable characteristic equation matrix be $S = MP$.

$$P = \left(M^T M\right)^{-1}\left(M^T S\right) \tag{26}$$

Rearranging the equations (23),(25) and (26), we obtain equation (27) as

$$\begin{bmatrix} 0 & 0 & 0.388 & 0 & 0 & 0.257 \\ 0.388 & 0 & 0.0147 & 0.257 & 0 & 0 \\ 0.0147 & 0.388 & -2.52 & 0 & 0.257 & 0 \\ -2.52 & 0.0147 & 0 & 0 & 0 & 0 \\ 0 & -2.52 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} K_{p1} \\ K_{i1} \\ K_{d1} \\ K_{p2} \\ K_{i2} \\ K_{d2} \end{bmatrix} = \begin{bmatrix} p_1 - 0.0581 \\ p_2 + 6.63 \\ p_3 + 0.126 \\ p_4 \\ p_5 \end{bmatrix} \tag{27}$$

The net characteristic equation of the stable closed loop system after solving equation (27), we obtain

$$s^5 + 50.1s^4 + 229s^3 + 362s^2 + 1399s + 201 = 0 \tag{28}$$

By comparing the equations (28) and (23), the PID gain values are obtained as:

$$K_{p1} = -520; K_{i1} = -79.7; K_{d1} = -78; K_{p2} = 1711; K_{i2} = 794; \\ K_{d2} = 312 \quad .$$

IV. Design of LQR controllers for cart inverted pendulum system

To design a full-state feedback regulator [7] for the plant described in equation (12) such that the control input vector is chosen as equation (29)

$$u(t) = -K\,x(t) \tag{29}$$

For an infinite final time, the quadratic objective function is given in equation (30).

$$J_\infty = \int_{t_o}^{\infty} [x^T Q x + u^T R u]\,dt \tag{30}$$

Here $u^T Ru$ is control energy in quadratic form, R is a cost matrix which puts a penalty on the control input, $x^T Qx$ is transient energy and Q is a state weighting matrix, which, puts a penalty on the state matrix. Both matrices Q and R are diagonal matrices whose entries put a penalty on respective states and input energy. The feedback gain matrix (K) is solved for the scalar objective function (J) by substituting equation (29) into (12) to obtain equation (31)

$$\dot{x} = \left[A - BK\right]x(t) = \left[A_{Closedloop}\right]x(t) \tag{31}$$

The optimal feedback gain matrix (K) in equation (32) is calculated by using MATLAB command, $K = lqr(A, B, Q, R)$.

$$K = \begin{bmatrix} -5 & -9.22 & 107 & 41.6 \end{bmatrix} \tag{32}$$

The diagonal entries of matrix Q are chosen by considering the range of each state and taking the inverse square. The position is bounded within 5cm, while the angle is bounded by 0.1 rad, thus giving the respective entries as 25 and 25 respectively. Derivatives of position and angle function in the Q matrix are penalized by unity. As the system only has a single input, the R matrix will only contain a single entry which is taken as 0.1 by considering the inverse square of the range of dc motor voltage. Closed-loop responses are presented with $R = 0.1$, and $Q = diag(1, 25, 1, 25)$. Controllability and observability matrices with full rank are given in equations (33) and (34), which implies that all states are simultaneously controllable and an observer design is possible.

$$C = \begin{bmatrix} 0 & 0.377 & -0.0071 & 0.2003 \\ 0.377 & -0.0071 & 0.2003 & -0.0075 \\ 0 & 0.367 & -0.0068 & 3.6885 \\ 0.3670 & -0.0068 & 3.6885 & -0.0718 \end{bmatrix} \tag{33}$$

$$O = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -0.019 & 0.2026 & 0 \\ 0 & -0.019 & 6.43 & 0 \\ 0 & 0.0004 & -0.002 & 0 \\ 0 & 0.0002 & -0.02 & 6.4 \end{bmatrix} \tag{34}$$

Using equation (32), eigenvalues of a matrix $(A - BK)$ are evaluated which gives the four poles of the desired closed loop system as in equation (35)

$$M = \begin{bmatrix} -2.585 + i0.206 \\ -2.585 - i0.206 \\ -0.992 + i0.942 \\ -0.992 - i0.942 \end{bmatrix} \tag{35}$$

For appropriate controller gain values, the fifth pole is now placed at -10. The characteristic equation derived from newly obtained poles is

$$s^5 + 17.2s^4 + 92.2s^3 + 232s^2 + 279s + 157 = 0 \tag{36}$$

By comparing the coefficients of equation (35) with (23), the PID values are obtained as $K_{p1} = -111$, $K_{i1} = -62$, $K_{d1} = -44$, $K_{p2} = 555$, $K_{i2} = 563$ and $K_{d2} = 135$.

It is observed that the values obtained of controller gain values are considerably low as compared to what we got using frequency response tuning, implying that the LQR method gives the minimum energy gain values for the PID controllers.

V. Comparative analysis and results

A. PID tuning using frequency response

The experimental response of cart and pendulum is shown in Fig.5 and Fig.6 respectively. The cart inverted pendulum system stabilized by tuning of PID controllers with frequency response, it is observed that the cart and pendulums are stabilized after 63 seconds but with more noise.

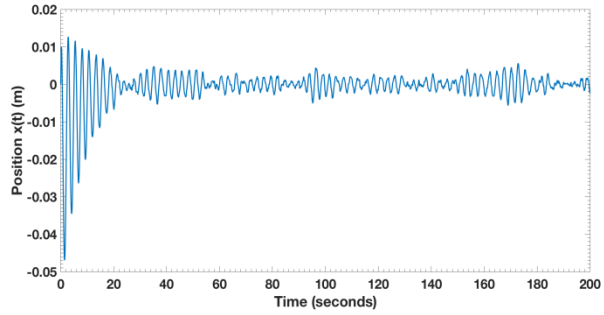


Fig.5 Experimental response of cart position

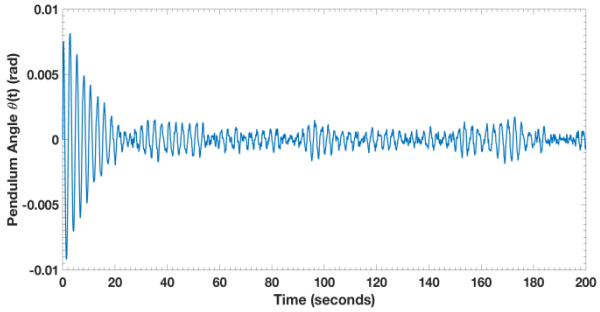


Fig.6 Experimental response of pendulum angle

B. PID tuning using LQR method

The experimental response of cart and pendulum is shown in Fig.7 and Fig.8 respectively.

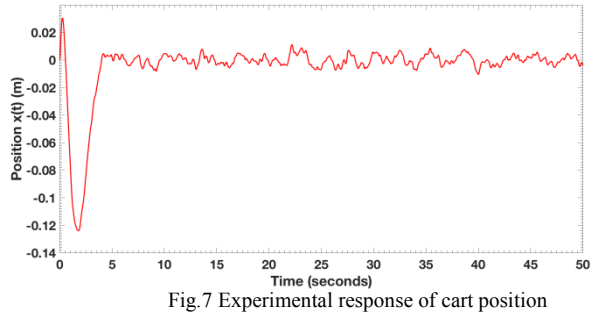


Fig.7 Experimental response of cart position

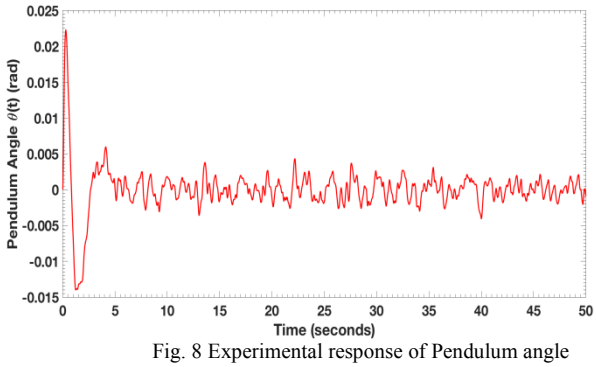


Fig. 8 Experimental response of Pendulum angle

The cart inverted pendulum system stabilized by tuning of PID controllers with LQR method, it is observed that the cart and pendulums are stabilized after 6 seconds. The comparative results of PID tuning with frequency response and LQR methods are given in Table.II.

TABLE II. COMPARATIVE RESULTS

Performance Parameters	Frequency Response		LQR	
	x(t)	Θ(t)	x(t)	Θ(t)
Settling Time (sec)	56.5	62.9	5.9	5.5
Overshoot (%)	13.21	8.45	23.48	18.28
Gain Margin (dB)	28.5	9.49	36.6	61.3
Phase Margin (dB)	82	68.7	50.2	62
Gain Crossover frequency (rad/sec)	3.94	1.41	0.789	0.791
Phase Crossover frequency (rad/sec)	4.876	2.43	1.21	1.35

VI. CONCLUSION

The stabilization of the Cart inverted pendulum system by two loops PID controllers tuned using frequency response and Linear Quadratic Regulator. While all design methodologies met the goal, the Linear Quadratic methods certainly lead to a better designed closed loop system. All methods lead to satisfactory transient response, but based on the stability margins and settling time, a system designed by frequency response gives an inferior closed loop system. Between these methods, the LQR gives the minimum energy gain values for the PID controllers, making it practical for real-time implementation.

ACKNOWLEDGMENT

The authors would like to thank Govt. of India and Govt. of NCT Delhi for providing funding through Delhi Technological University under TEQIP-II scheme to procure Digital Pendulum System experimental setup.

REFERENCES

[1] E.J.Davison, "Benchmark problems for the control system design: report of the IFAC Theory Committee," Schlossplatz 12, A-2361 Laxenburg, Austria, 1990.

[2] Jonathan P. How, "Benchmarks," *IEEE Control Systems Magazine*, pp. 6–7, 2015.

[3] FeedbackInstruments, "Digital Pendulum Control Experiments Manual," 2016.

[4] Sudarshan K. Valluru, Madhusudan Singh, Bharat Bhushan, "Comparative Analysis of PID, NARMA L-2 and PSO Tuned PID Controllers for Nonlinear Dynamical System," *J. Autom. Syst. Eng.*, vol. 9, no. 2, pp. 94–108, 2015.

[5] Sudarshan K. Valluru and Madhusudan Singh, "Stabilization of Nonlinear Inverted Pendulum System Using MOGA and APSO Tuned Nonlinear PID Controller," *Cogent Eng.*, vol. 4, no. 1, pp. 1–15, 2017.

[6] B. Ghosh, A., Krishnan, T.R., Subudhi, "Robust proportional – integral – derivative compensation of an inverted cart-pendulum system : an experimental study," *IET Control Theory Appl.*, vol. 6, no. 8, pp. 1145–1152, 2012.

[7] J. B. He, Q. G. Wang, and T. H. Lee, "PI/PID controller tuning via LQR approach," *Chem. Eng. Sci.*, vol. 55, pp. 2429–2439, 2000.

View publication stats

Computational Materials Science 151 (2018) 65–72

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Few-layer p-type phosphorene sheet: An efficient transparent conducting electrode in silicon heterojunction solar cell

Chandra Kamal Borah[a], Pawan K. Tyagi[b,*], Sanjeev Kumar[a], Kamlesh Patel[a,c]

[a] Department of Physics, Rajiv Gandhi University, Arunachal Pradesh 791112, India
[b] Department of Applied Physics, Delhi Technological University, New Delhi 110042, India
[c] Department of Electronic Science, University of Delhi South Campus, New Delhi 110021, India

ARTICLE INFO

Keywords:
Phosphorene
Silicon heterojunction solar cell
AFORS-HET software
Transparent conducting electrode

ABSTRACT

In this article, a structure of heterojunction solar cell configured as Ti (front contact)/p-phosphorene/n-cSi/Ag (back contact) has been simulated by using AFOSR-HET v2.5 software. Here, electronic nature of phosphorene layer which is anisotropic along armchair and zigzag directions has been considered as 3D in nature instead of 2D. After optimizing the parameters of p-type single layer phosphorene and n-type crystalline silicon (n-cSi) maximum power conversion efficiency 15.51% and 14.06% has been achieved for armchair and zigzag phosphorene, respectively. The temperature dependence of the cell performance has been studied and maximum efficiency has been found at 300 K for both armchair and zigzag phosphorene layer. Further, we have studied the effect of metal contact on the cell performance. We have also simulated the solar cell taking experimentally available parameters of n-cSi and efficiencies of 13.29% and 14.06% are found in cases of the armchair and zigzag phosphorene layers, respectively. Finally, we have demonstrated that multilayer p-type phosphorene layer could act as an effective transparent conducting electrode in the silicon heterojunction solar cell.

1. Introduction

Recently, as a physical analogue to the graphene, single atomic thick layer of black phosphorus (BP) known as phosphorene has emerged as a promising material for next-generation transistor applications due to its superior carrier transport properties. Previously, BP's instability in the air was considered as a major challenge for its potential application in the devices that operate in an ambient environment [1]. Recently, air stability of BP layers has been improved by depositing an Al_2O_3 layer on the surface [1,2]. In further advances, it is reported that BP possesses a direct band gap of 0.3–2.0 eV depending on the layer number and mobility as high as 10,000 $cm^2/V\,s$ compared to 200–517 $cm^2/V\,s$ in other 2D materials [3–5]. Recently Ge et al. [6] reported that as-synthesized BP is p-type with carrier density $4.2 \times 10^{18}/cm^3$, equivalent to $2 \times 10^{11}/cm^2$ for each layer of BP. Du et al. [7] reported that chemical doping with 2,3,5,6-tetrafluoro-7,7,8,8-tetracyanoquinodimethane (F4-TCNQ) makes the few-layer black phosphorous (FLBP) to p-type. Also, He et al. [8] theoretically predicted that F4-TCNQ enhances the p-type conductivity in single layer phosphorene due to strong charge transfer between its molecule and layer. Furthermore, the sheet resistance of mechanically exfoliated layer has been reported to be in the range of 100–1000 kΩ/sq and after doping with F4-TCNQ this reduces to 2.5 times [1,7]. Optical absorbance was reported to be 2.8% and 5.5% in single and bilayer phosphorene, respectively confirms its high transparency [9]. The work function of BP decreased rapidly from 5.16 eV to 4.5 eV as layer number increased from 1 to 5 [10]. Similarly the dielectric constant also increased rapidly from 2.9 (bilayer) to 3.5 (trilayer) [11]. Above discussed properties of FLBP motivates us to use the p-type FLBP in silicon-based heterojunction solar cell as a transparent conducting electrode. Generally, a transparent conducting electrode in the heterojunction solar cell allows the light to pass through the interface and also introduces a built-in potential at phosphorene/Si interface to separate the photogenerated carriers. Recently, application of phosphorene in various areas such as transistors, memory devices, and batteries has been demonstrated [2]. However, the utilization of phosphorene in a solar cell has not been explored much and few reports on the application of phosphorene in solar cell application are available. In 2016, theoretically, Hu et al. [12] proposed that edge-modified phosphorene nanoflakes can act as donor and acceptor materials in heterojunction solar cell and efficiency of 20% has been predicted. Experimentally, liquid phase exfoliated black phosphorous has been used in the organic solar cell. Lin et al. [13] have used the thin layer of phosphorene as anode buffer layer in device structure configured as Glass/ITO/PEDOT:PSS/

* Corresponding author.
E-mail addresses: pawankumartyagi@gmail.com, pawantyagi@dtu.ac.in (P.K. Tyagi), sanjeev.kumar@rgu.ac.in (S. Kumar).

https://doi.org/10.1016/j.commatsci.2018.04.059
Received 10 January 2018; Received in revised form 27 March 2018; Accepted 24 April 2018
0927-0256/ © 2018 Elsevier B.V. All rights reserved.

BP/PTB7:$PC_{71}BM$/Ca/Al. After inserting phosphorene layer, a reduction in efficiency from 7.72% to 6.52% was reported. Thus, optimization of layers' parameters of phosphorene as well as silicon is required to obtain a cost-effective high-efficiency solar cell. Hence in this report, the cell structure of p-phosphorene/n-cSi under air mass 1.5 (AM1.5) illumination with a power density of 100 W/cm^2 using AFORS-HET (Automat FOR Simulation of HETero-structures) software was simulated. This software solves the 1D semiconductor equations-Poisson's equation and the transport equation for electrons and holes with the help of finite difference under different conditions, i.e. steady state condition, equilibrium condition, etc. In the proposed simulation, we have taken the p-phosphorene as a conducting electrode and n-cSi as an absorber layer. Furthermore, we have considered p-phosphorene 3D in electronic nature instead of 2D. The parameters of both p-phosphorene and n-cSi layers are optimized to obtain the highest efficiency for p-phosphorene/n-cSi solar cell. The simulation was also performed considering the anisotropic nature of phosphorene layer. The objective of the simulation is to study the independent effects of anisotropic behavior of phosphorene as well as Si layer on the performance of the cell and understand the insight behind cell response.

2. Simulation structure

As shown in Fig. 1(a), the structure of heterojunction solar cell configured as Al_2O_3/Ti/p-phosphorene/n-cSi/Ag has been simulated. In order to prevent the degradation of phosphorene layer in air, capping layer of Al_2O_3 is employed [1,2] and to minimize the reflection from the layer, its thickness was kept 40 nm [14]. Here Ti [2,15–17] and Ag are used as front and back contacts, respectively. Since Lambert-Beer model is applied, which is an in-built feature of the AFORS-HET software, the

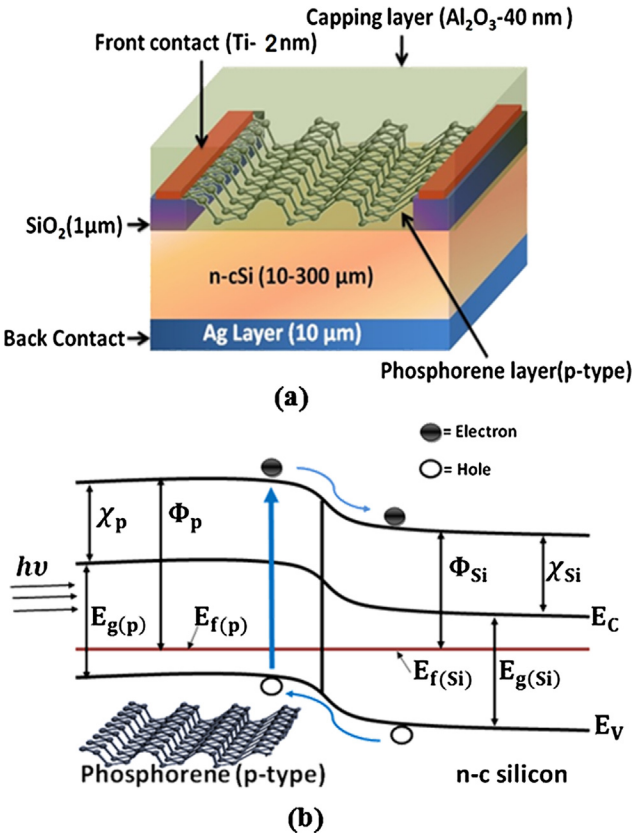


Fig. 1. Phosphorene-silicon heterojunction solar cell. (a) Schematic diagram of proposed p-phosphorene/n-cSi heterojunction (b) Schematic of band diagram after the formation of heterojunction displaying the direction of electron and hole transport.

Table 1
Details of the front and back contact parameters used in simulation.

Contact parameters	Front Contact	Back Contact
Contact Material	Titanium (Ti)	Silver (Ag)
Width (cm)	2×10^{-8}	0.001
[n, k] values	Ti. nk	Ag. nk
Metal work function (eV)	4.33	4.26
Absorption loss	Ti.abs	0
External reflection constant	Ti.ref	0
Surface condition	Plane	Plane
Internal reflection constant	0	0

reflection from the phosphorene/Si interface is not taken into account [18]. The [nk] file for Ti has been created using the data reported by Rakić et al. [19]. In addition, the reflectance file of Ti has been made using FILMETRIC reflectance calculator software based on the complex matrix form of the Fresnel equation [20]. Absorbance file of Ti has been created using the relations αx and $\alpha = \frac{4\pi k}{\lambda}$, here where x, k and λ is the thickness, extinction coefficient and wavelength, respectively. The in-built file of [nk] for Ag was used for simulation. Details of the front and back contact parameters have been tabulated in Table 1. In simulated structure, the transport of minority carriers was taken along the c-axis, therefore insulating SiO_2 layer was introduced to stop the direct transport of charge carriers between Ti and n-cSi. The Fermi wavelength of the 2D system is given by the relation $\lambda_F = \sqrt{\frac{2\pi}{n_{2D}}}$, where n_{2D} is the carrier density in cm^{-2} [21]. Recently, Saito et al. have reported the sheet carrier density in BP up to $3 \times 10^{13}\,cm^{-2}$ [22]. So, if n_{2D} approaches $\sim 3 \times 10^{13}\,cm^{-2}$, then λ_F is found to be 4.576 nm. Hence, doped six layer phosphorene (~3.96 nm) will satisfy the condition $\langle t \rangle \leq \lambda_F$, where $\langle t \rangle$ is the average thickness for 2D electron system. Therefore, the facts discussed above motivated us to consider multilayer phosphorene as quasi 3D in nature. Here, the quantum confinement was achieved by reducing the thickness. Thickness of single layer phosphorene is fixed at 0.66 nm [10] and [nk] file has been created on the basis of reported values of n and k [23]. Due to anisotropic nature, properties like effective mass and mobility of carriers are different along armchair and zigzag direction [4,10].Thus, the simulation was first carried out considering armchair and then extended for zigzag separately. The variable parameters i.e., acceptor concentration (N_A), effective conduction band and valance band density (N_C/N_V), band gap (E_g), electron affinity (χ), dielectric constant (ε_r), electron mobility (μ_n), hole mobility (μ_p) and layer thickness are optimized. Other values as well as range of the parameters of phosphorene and silicon have been tabulated in Table 2. The values of N_C/N_V for phosphorene have been estimated using following equation:

$$N_{C/V} = 2\left(\frac{2\pi k_B T m^*_{e/h}}{h^2}\right)^{\frac{3}{2}} \quad (1)$$

Here k_B = Boltzmann constant, T = temperature in Kelvin, m^*_h = effective mass of electrons, m^*_e = effective mass of holes, h = Planck's constant [24]. The thermal velocity of carriers has been determined using the relation $V_{th} = \sqrt{\frac{3k_BT}{m^*_{e/h}}}$ [25]. In a second step we have further optimized parameters of n-cSi by varying the donor concentration (N_D), conduction and valance band density (N_C/N_V), electron affinity (χ) and thickness. In this simulation work, the range of parameters of p-phosphorene are kept to the reported values in the previous literatures [4,6,10,11,26–31]. Similarly, the ranges of parameters of n-cSi layers are taken from the available literatures [32–37]. Here we have considered the ranges of the parameters for single crystal Si (1 0 0) [32,33,38] and we have not taken into account the lattice matching between layers. The details of simulation parameters of p-phosphorene and n-cSi are given in Table 2.

Schematic of energy band diagram of p-phosphorene/n-cSi interface has been shown in Fig. 1(b). At the interface, electrons are diffused

Table 2
Ranges of the parameters taken for simulation in the proposed solar cell.

Input parameters	Phosphorene (Arm chair)	n-cSi
Thickness (cm)	6.6×10^{-8}–1.32×10^{-6}	0.001–0.03
Dielectric constant (ε_r)	2.8–10.2	11.9
Electron affinity χ (eV)	3.6–4.1	4.02–4.22
Band gap, E_g (eV)	0.30–1.50	1.12
Effective conduction band density, N_C (cm^{-3})	1×10^{12}–2×10^{19}	3×10^{13}–3×10^{19}
Effective valence band density, N_V (cm^{-3})	1×10^{12}–2×10^{19}	3×10^{13}–3×10^{19}
Effective electron mobility, μ_n ($cm^2/V\,s$)	1.58×10^{-2}–1.58×10^{3}	1500
Effective hole mobility, μ_p ($cm^2/V\,s$)	6.4×10^{-2}–6.4×10^{3}	500
Acceptor concentration N_A (cm^{-3})	1×10^{10}–1×10^{18}	0
Donor concentration N_D (cm^{-3})	0	1×10^{12}–1×10^{17}
Electron thermal velocity, V_e ($cm\,s^{-1}$)	3×10^{7}	1×10^{7}
Hole thermal velocity, V_h ($cm\,s^{-1}$)	3.1×10^{7}	1×10^{7}
Layer density, ($g\,cm^{-1}$)	2.69	2.328
Auger electron recombination coefficient, ($cm^6\,s^{-1}$)	$\sim 4.098 \times 10^{-43}$	2.8×10^{-31}
Auger hole recombination coefficient, ($cm^6\,s^{-1}$)	$\sim 4.098 \times 10^{-43}$	1×10^{-31}
Band-to-band recombination coefficient, ($cm^3\,s^{-1}$)	0	0
Defect properties	Single Acceptor	Single Acceptor
Total trap density (cm^{-3})	1.36×10^{15}	1×10^{14}
Characteristic energy (eV)	0.0216	0.56
Effective mass of electron (m^*/m_0)	0.17	0.98
Effective mass of hole (m^*/m_0)	0.15	0.49

from n-cSi to p-phosphorene and holes from p-phosphorene to n-cSi until thermal equilibrium gets established. During this process, first depletion region forms and then due to the built-in electric field, minority carriers drift in opposite direction. This process continues until the drift current balances the diffusion current, thereby thermal equilibrium is obtained and bending of the energy band near Fermi level occurs. This built-in potential (eV_{bi}) is also responsible for charge separation in the solar cell. The built-in potential (eV_{bi}) in p-phosphorene and n-cSi can be defined in terms of their work function by the following equations:

$$\begin{aligned} eV_{bi} &= \Phi_p - \Phi_{Si} \\ &= \chi_p + \left[E_{g(p)} - \left\{ -k_B T \ln\left(\frac{N_A}{N_V}\right)_{(p)} \right\} \right] - \left[\chi_{Si} + \left\{ -k_B T \ln\left(\frac{N_D}{N_C}\right)_{(Si)} \right\} \right] \\ &= E_{g(p)} + (\chi_p - \chi_{Si}) + k_B T \ln\left\{ \left(\frac{N_A}{N_V}\right)_{(p)} \left(\frac{N_D}{N_C}\right)_{Si} \right\} \end{aligned} \tag{2}$$

Here Φ_p and Φ_{Si} are the work function of p-phosphorene and n-cSi respectively, $E_{g\,(p)}$ is the band gap energy of p-phosphorene, χ_P and χ_{Si} are the electron affinity of p-phosphorene and n-cSi respectively, k_B is the Boltzmann constant, N_A and N_D are acceptor and donor concentrations, N_C and N_V are the conduction band density and the valence band density of n-cSi and p-phosphorene, respectively [39].

As p-phosphorene/n-cSi junction is illuminated, electron-hole pairs (EHP) are generated and then separated due to built-in potential. These free minority carriers move towards the p-phosphorene and n-cSi, respectively, as shown in Fig. 1(b). Under open circuit condition, these photo-generated carriers enhance the carrier density in the respective regions. Due to the accumulation of these carriers, Fermi level shifts upward in n-cSi and downward in p-phosphorene, respectively and as a consequence V_{OC} develops. Similarly, during short-circuiting photo-generated charge carriers do not accumulate. This happens due to their collocation of carriers at the contact electrode, which results in short circuit current (I_{SC}) in the external circuit.

3. Results and discussions

3.1. Optimization and simulation of the p-phosphorene layer (Arm chair direction)

In this section, first, we have optimized acceptor concentration (N_A) of p-phosphorene layer in the range from 10^{10} to $10^{18}\,cm^{-3}$ and results were plotted in Fig. 2(a and b). As shown in Fig. 2b, V_{OC} remained constant at 516.4 mV until N_A increased up to $10^{14}\,cm^{-3}$ and then suddenly decreased to 514.8 mV at $N_A = 10^{15}\,cm^{-3}$ and remained constant till N_A reached to $10^{18}\,cm^{-3}$. Similarly, the short circuit current density (J_{SC}) remained constant for N_A up to $10^{14}\,cm^{-3}$ and then J_{SC} decreased from 26.37 to 24.3 mA/cm^2. After $N_A = 10^{15}\,cm^{-3}$, decrement in J_{SC} became slow. It is reported that at low N_A, the mobility of photogenerated carriers remained nearly constant and limited by scattering with phonon while at higher N_A mobility was decreased due to the scattering with ionized doping atoms [40]. It is known that carrier's lifetime depends on mobility given by the relation $\mu_i = \frac{e}{m^*}\tau_i$. Furthermore, as μ_i gets reduced, diffusion coefficient (D_i) also reduced. The diffusion coefficient, according to the Einstein relation is given as $\frac{D_i}{\mu_i} = \frac{k_B}{e}T$ [24]. Since at higher N_A both τ_i and D_i reduced, consequently diffusion length ($L_i = \sqrt{D_i\tau_i}$) reduced. Thus, J_{SC} decreased due to the reduction of L_i following the relation:

$$J_{SC} = qG(L_n + L_p) \tag{3}$$

Here L_n and L_p is the diffusion length for electron and hole, respectively and G is carrier generation rate and q is the electric charge of carriers [41]. Further, V_{OC} was found to be constant at 516.4 mV till N_A reached to $10^{14}\,cm^{-3}$ and then slightly decreased and became constant at 514.8 mV for higher N_A. It is reported that V_{OC} does not depend on N_A if condition $\Delta n_o \gg N_A$ is satisfied, here Δn_o is the excess electron-hole pairs [35]. Thus V_{OC} remained constant till $N_A = 10^{14}\,cm^{-3}$ and is given by the equation:

$$V_{OC} = \frac{nqT}{k}\ln\left(\frac{J_{SC}}{J_0} + 1\right) \tag{4}$$

Here n, J_o are ideality factor and reverse saturation current density (mA/cm^2), respectively [41]. J_o is defined as:

$$J_o = eA\left(\frac{D_p P_n}{L_n} + \frac{D_n n_p}{L_p}\right) \tag{5}$$

where A is the area of the junction [24]. At N_A above $10^{14}\,cm^{-3}$, doping may reduce L_i and results in the enhancement of J_o following Eq. (5). This increment in J_o results in a decrement in V_{OC} within the range 10^{14} to $10^{15}\,cm^{-3}$. Further, for N_A above $10^{15}\,cm^{-3}$ J_o increased whereas J_{SC} decreased. This resulted in a reduction of V_{OC}. But at the same time, N_A will increase V_{bi} according to Eq. (2), which leads to increases in V_{OC}. These two processes might give a constant value in V_{OC} after $N_A = 10^{15}\,cm^{-3}$. From Fig. 2(a) it was observed that the fill factor (FF) and efficiency (η) remain constant up to the value of $N_A = 10^{14}\,cm^{-3}$ and $N_A = 10^{13}\,cm^{-3}$, respectively due to the constant value of V_{OC} and J_{SC}. After $N_A = 10^{14}\,cm^{-3}$, there is a sudden increase of FF. This sudden increase may be due to the decrease of V_{OC} and J_{SC} and sheet resistance which reduced due to doping in this range [42]. After $N_A = 10^{15}\,cm^{-3}$, the FF decreases slowly. The maximum value of η was found to be 7.493% at $N_A = 10^{14}\,cm^{-3}$.

After optimizing N_A to $10^{14}\,cm^{-3}$, the value of N_C/N_V was optimized within the range 1×10^{15} to $2 \times 10^{18}\,cm^{-3}$ as shown in Fig. 2(c and d). It was observed that both V_{OC} and J_{SC} increased linearly with N_C/N_V. But, according to Eq. (2) both should decrease with the decrease in V_{bi}. In the present case, both V_{OC} and J_{SC} were found to increase as shown in Fig. 2(d). Increase of V_{OC} can be explained in terms of carrier

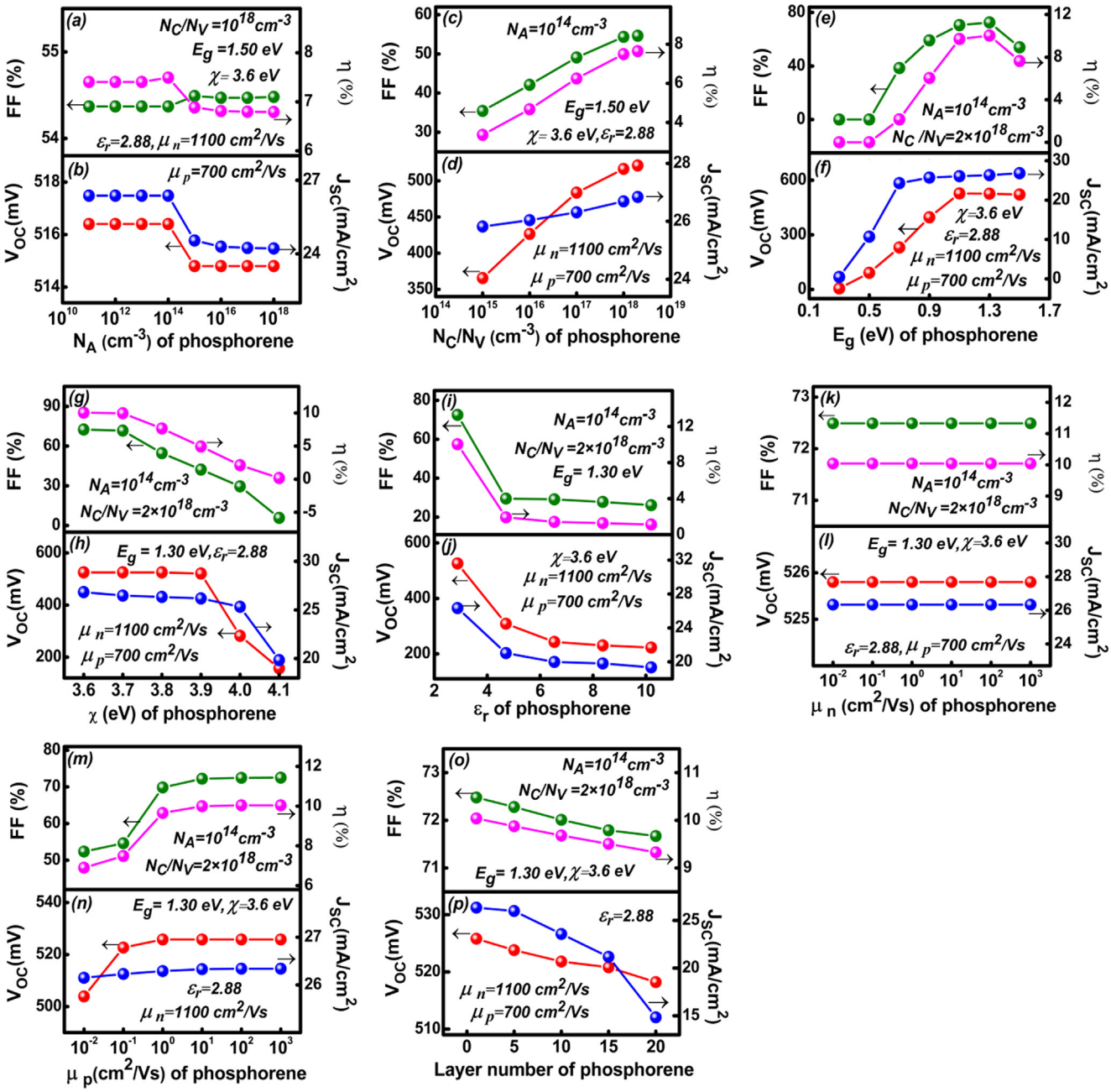


Fig. 2. Optimization of p type armchair phosphorene layer's parameters, (a-b) donor concentration N_D (cm^{-3}), (c-d) effective conduction band and valance band density N_C/N_V (cm^{-3}), (e-f) the band gap energy E_g (eV), (g-h) electron affinity χ (eV), (i-j) dielectric constant ε_r, (k-l) electron mobility μ_n (cm^2/Vs), (m-n) hole mobility μ_p (cm^2/Vs), (o-p) layer number of phosphorene where n-cSi parameters are kept as, thickness = 220 μm, $N_D = 9 \times 10^{14}\,cm^{-3}$, $N_C/N_V = 1 \times 10^{19}\,cm^{-3}$ and χ = 4.22 eV. The other parameters are given in the Table 2.

concentration as $V_{OC} = \frac{kT}{q}\ln\left(\frac{n.p}{n_i^2}\right)$ here n_i, n and p are intrinsic, electron and hole concentration, respectively. The n and p can also be written as $n = n_o + \Delta n$ and $p = p_o + \Delta p$, where n_o and p_o are concentration of electron and hole in equilibrium. The Δn and Δp are the excess photogenerated electron and hole. The intrinsic carrier concentration is related with N_C and N_V as $n_i = (N_C N_V)^{\frac{1}{2}} e^{\frac{-E_g}{2kT}}$. Thus in this case, $n_o p_o = {n_i}^2$, $p_o = N_A$ and $\Delta n = \Delta p$, so V_{OC} can be written as $V_{OC} = \frac{kT}{q}\ln\left(1 + \frac{\Delta n}{N_A} + \frac{\Delta n^2}{n_i^2} + \frac{\Delta n N_A}{n_i^2}\right)$ [43]. Therefore, we can assume that V_{OC} was increased with N_C/N_V due to an increase of n_i. In addition, high N_C may lead to high absorption and enhanced J_{SC} [44]. The FF and η also found to be increased with N_C. The maximum efficiency η = 7.645% has been achieved at $N_C/N_V = 2 \times 10^{18}\,cm^{-3}$.

After optimizing N_A ($10^{14}\,cm^{-3}$) and N_C/N_V ($2 \times 10^{18}\,cm^{-3}$) we have optimized band gap (E_g) in the range 0.30–1.50 eV and results were plotted in Fig. 2(e and f). It is reported that the band gap of phosphorene layer decreased from 1.50 eV to 0.595 eV as layer number decreased from 1 to 5 [1,4,10]. We observed that as band gap increased up to 1.30 eV, V_{OC} as well as J_{SC} increased linearly up to 527.3 eV and 26.34 mA/cm^2, respectively. This variation in V_{OC} and J_{SC} can be understood from Eqs. (2)–(4). As E_g increases, V_{bi} also increases {Eq. (2)} and high V_{OC} obtains. The increment in V_{bi}, separate the photogenerated carriers effectively and enhances J_{SC}. Above E_g = 1.30 eV, V_{OC} decreased to 521.1 eV and J_{SC} slightly increased to 26.84 mA/cm^2. This decrease in V_{OC} and increase in J_{SC} might happen due to the fact that the doping density $N_A = 10^{14}\,cm^{-3}$ shrinks E_g to 1.30 eV. In Fig. 2(e). FF increases from 0% to 72.49% as E_g increases to 1.30 eV.

This increment might be due to the decrease of J_o. Beyond $E_g = 1.30$ eV, FF decreases to 54%. The decrement in E_g enhances intrinsic carrier concentration (n_i) as well as J_o and results in a decrease of FF. Similarly, η founds to increase as E_g increases to 1.30 eV and further increase of E_g results the decrease in η. The decrease in η might be due to the decrease in V_{OC} and FF.

The optimization of χ_p was done in the range of 3.6–4.1 eV as shown in Fig. 2 (g-h). It was observed that as electron affinity increased, first V_{OC} remained constant at 525.8 mV and then decreased abruptly for $\chi_p > 3.8$ eV and reached to 157 mV at $\chi_p \approx 4.1$ eV. Similarly, J_{SC} was found to reduce slightly from 26.34 to 25.19 mA/cm^2 as χ_p increased to 3.9 eV and then suddenly decreased to 19.84 mA/cm^2 for $\chi_p = 4.1$ eV. Considering the band diagram (Fig. 1b) and Eq. (2), these variations can be understood as follows: as χ_p increased bottom edge of conduction band bend downward as a consequence $E_{g(p)}$ decreased. This finally lowered the built-in potential which reduced V_{OC}. This lowering of V_{bi} decreased the separation of photogenerated carriers to their respective contacts. Thus, J_{SC} also decreased with the increase of χ_p. The decrease of FF might be due to recombination of photogenerated carriers and the maximum value of efficiency $\eta = 10.04\%$ was achieved at $\chi_p = 3.6$ eV.

The dielectric constant (ε_r) has been reported to increase with layer numbers and has been found to be 2.9, 3.5 and 10.2 for bilayer, trilayer and bulk BP, respectively [11]. Variations of cell parameters with ε_r have been studied and plotted in Fig. 2(i and j). V_{OC} decreased from 525.8 mV to 222.7 mV as ε_r increased from 2.8 to 10.2. J_{SC} was also found to decrease from 26.34 mA/cm^2 to 19.35 mA/cm^2. Since V_{bi} and ε_r are inversely proportional and related as $V_{bi} \propto \frac{1}{\varepsilon_r}$ [24,43], therefore as ε_r increased V_{bi} decreased and resulted in the decrease of V_{OC} as well as J_{SC}. FF and η was found to decrease with increase of ε_r. The highest efficiency $\eta = 10.04\%$ was achieved at $\varepsilon_r = 2.88$. The efficiency and other parameters of the cell are found to be independent of electron mobility (μ_n) of the p-phosphorene layer.

As reported in Refs. [4,8], the mobility of BP is hole-dominated and reported to be in the order of 10,000 cm^2/Vs (in a single layer). Thus, we also studied the response of the cell parameters with hole mobility. As shown in Fig. 2(m and n), as hole mobility increased V_{OC} increased from 503.9 mV to 525.8 mV and J_{SC} from 26.15 mA/cm^2 to 26.34 mA/cm^2. If mobility will increase then recombination of charge carrier will decrease and τ_p, D_p and L_p will increase. Thus V_{OC} enhanced due to J_o whereas J_{SC} due to the increase of L_p. Similarly, FF was found to increase from 52.37% to 72.51%, which might be due to the decrease of J_o. The maximum $\eta = 10.04\%$ was achieved at even mobility $\mu_p = 640$ cm^2/V s.

After optimizing above-mentioned parameters the simulation was performed by varying the layer number from 1 to 20. It was observed that as the layer number increased V_{OC} slightly decreased from 525.8 mV to 518.2 mV, whereas J_{SC} remarkably decreased from 26.34 mA/cm^2 to 14.83 mA/cm^2. This might have happened due to lowering of electron-hole pair generation rate which depends on the transmittance of phosphorene. It is reported that as the layer number increased, the transmittance was reduced which resulted in reduced electron-hole pair generation rate. Due to lower V_{OC}, J_{SC} and FF consequently η was found to decrease sharply and η was decreased from 10.04% to 9.328% as layer number increased from 1 to 20. The FF was reduced due to lowering of sheet resistance which reduced as layer number increased.

3.2. Optimization and simulation of n-cSi wafer

After optimizing the parameters of p-phosphorene, the further simulation was carried out to optimize the n-cSi parameters while keeping the parameters of p-phosphorene at their best-optimized values. As shown in Fig. 3(a and b) it was observed that as N_D was increased from 10^{12} cm^{-3} to 10^{13} cm^{-3}, V_{OC} increased very slightly from 293 mV to 294.1 mV. But for $N_D > 10^{13}$ cm^{-3} V_{OC} was found to increase linearly and reached up to 655.5 mV. Similarly J_{SC} increased slowly from 0.312 mA/cm^2 to 0.356 mA/cm^2 as N_D increased from 10^{12} cm^{-3} to 10^{13} cm^{-3}. After $N_D > 10^{13}$ cm^{-3}, J_{sc} was increased almost linearly and reached up to 28.66 mA/cm^2. Following Eq. (2) high N_D enhanced V_{bi}, which facilitates the movement of generated charge carriers towards their respective electrodes effectively as result V_{OC} and J_{SC} increased. FF and η were found to increase with N_D. The increment of FF can relate to decrease of the resistivity of the n-cSi substrate [42]. The maximum value of $\eta = 15.51\%$ was obtained at $N_D = 10^{17}$ cm^{-3}.

After optimizing N_D to 10^{17} cm^{-3} the simulation was further carried out to optimize N_C/N_V within the range of 1×10^{19}–3×10^{19} cm^{-3} and results have been shown in the Fig. 3(c and d). V_{OC} was found to decrease from 655.5 mV to 597.7 mV with the increase of N_C/N_V whereas J_{SC} almost remains constant at 28.66 mA/cm^2. According to Eq. (2), as N_C/N_V increases V_{bi} decreases, thus the trapping of charge carrier inside the silicon layer might be increased. Further increase in N_C/N_V can directly relate to effective mass (m^*) of charge carriers which leads to the decrease of mobility as well as L_i of charge carriers according to Eqs. (1), (3) and (4). So J_o increases according to Eq. (5). This decrease in V_{bi} and increase in J_o decreases the V_{OC}. It has been also reported that N_C/N_V influences the band gap as well as absorption of photons [44]. From Eq. (3), it can be estimated that at $N_C/N_V = 1 \times 10^{19}$ cm^{-3}, the high absorption of photons in active silicon layer increased generation rate of the electron-hole pair (G). However, the high value of G may be further balanced by the lower value of L_i consequently leading to a constant value of J_{SC}. FF and η were found to decrease linearly up to $N_C/N_V = 2 \times 10^{19}$ cm^{-3} and then FF slightly enhanced whereas η decreased. The initial decrement in FF and η may be due to an increase of J_o at $N_C/N_V > 2 \times 10^{19}$ cm^{-3}. At higher N_C/N_V, FF decreases sharply due to its dependence on V_{OC}. Similarly, a linear decrease in η is found to depend directly on FF. The highest value of $\eta = 15.51\%$ was obtained at $N_C/N_V = 10^{19}$ cm^{-3}.

After optimization of N_D and N_C/N_V, electron affinity (χ) was optimized within the range of 4.02–4.22 eV and results are shown in Fig. 3(e and f). V_{OC}, J_{SC} and FF were found to be constant at 653.9 mV, 8.69 mA/cm^2 and 82.50%, respectively. This can be attributed to the fact that χ does not affect the position of Fermi level, thus V_{bi} does not change. This makes V_{OC} constant. Slight variation in η was observed and the maximum value of 15.51% at $\chi = 4.22$ eV was obtained.

We have then studied the influence of wafer thickness after optimization of N_D, N_C/N_V, χ. The thickness was varied within the range 10 μm to 300 μm. The simulated results are shown in Fig. 3(g and h). We observed that V_{OC} as well as J_{SC} increases linearly as thickness increases from 10 μm to 50 μm. V_{OC} increased from 652.3 mV to 655.5 mV and J_{SC} increased from 20.85 mA/cm^2 to 28.64 mA/cm^2. For thick wafer i.e. > 50 μm, all the cell parameters became constant. The maximum efficiency $\eta = 15.37\%$ is achieved at wafer thickness 50 μm. As Si thickness increases, absorption of photon increases and results in the generation of more electron-hole pairs. Moreover, in thin wafer back contact is close to the depletion region and enhances the collection of photogenerated carriers [45]. Thus V_{OC} and J_{SC} might have improved due to more photon absorption. After thickness > 50 μm, the constant value of J_{SC} might be obtained due to balance between G and L_i. It has been reported earlier that the value of V_{OC} become independent of wafer thickness (W_b) if $L_i \ll W_b$ [46,47]. The maximum value $\eta = 15.51\%$ was achieved at 100 μm of the n-cSi wafer.

After optimizing Si parameters the structure was again simulated by varying the layer number of phosphorene from single layer to 20 layers. The results are shown in Fig. 3(i and j), η was found to decrease from 15.51% to 13.21% as layer number increased from 1 to 20. To have better practical realization, we have further simulated considering the parameters of commercially available Si. We have taken the value of $N_D = 10^{15}$cm^{-3}, $N_C = 2.8 \times 10^{19}$ cm^{-3}, $N_V = 1.04 \times 10^{19}$ cm^{-3} $\chi = 4.05$ eV, $\mu_n = 1500$ cm^2/Vs and $\mu_p = 450$ cm^2/Vs for n-cSi as reported in Refs. [37,48]. We have varied the thickness of Si wafer from

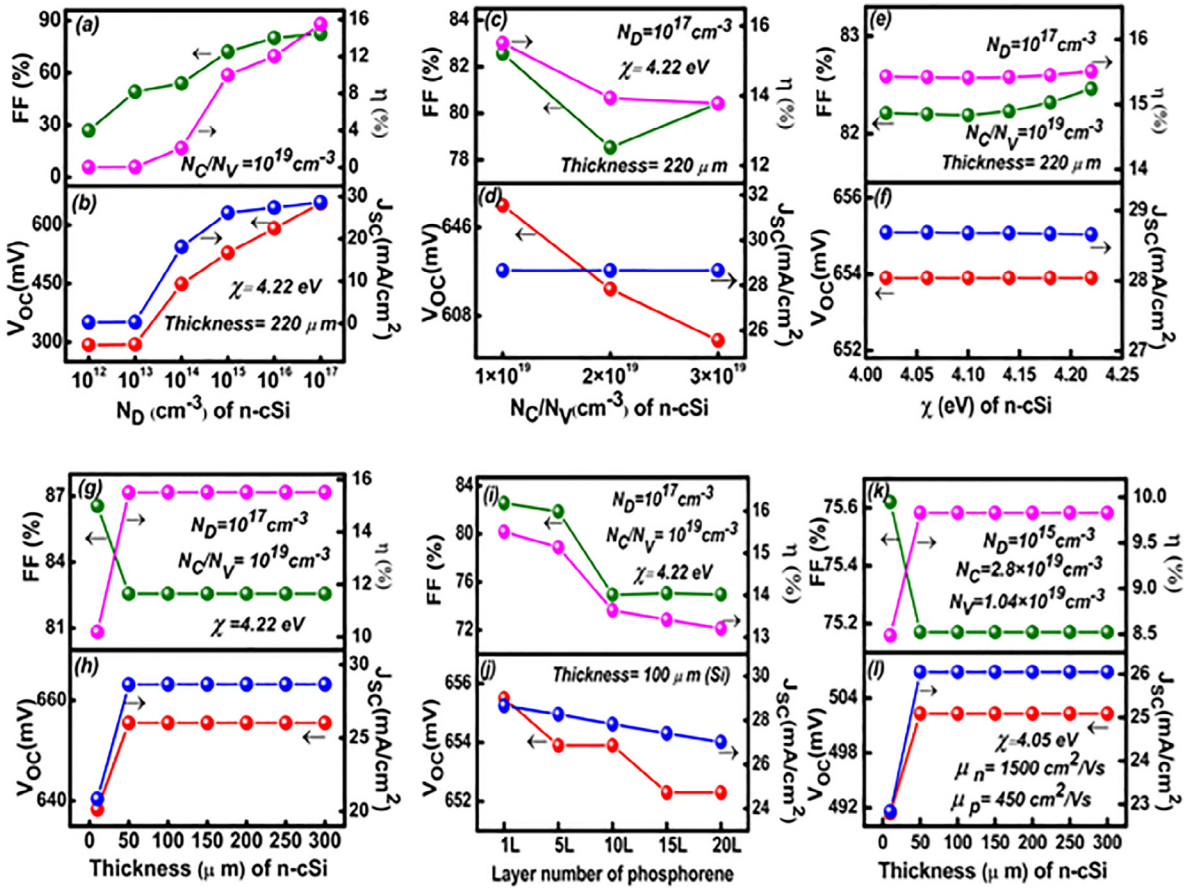


Fig. 3. Optimization of n-cSi parameters, (a-b) acceptor concentration N_A (cm^{-3}), (c-d) effective conduction band and valance band density N_C/N_V (cm^{-3}), (e-f) electron affinity χ(eV), (g-h) thickness (μm) where the optimized parameters for phosphorene (Arm-chair) are kept as, $N_A = 10^{14}$ cm^{-3}, $N_C/N_V = 2 \times 10^{18}\,cm^{-3}$,$E_g = 1.3$ eV, χ = 3.6 eV, $\mu_n = 1100\,cm^2/Vs$ and $\mu_p = 700\,cm^2/vs$. (i-j) Optimization of phosphorene layer number after optimizing the n-cSi wafer thickness. (k-l) Optimization of the thickness of commercially available silicon wafer. The parameters of phosphorene are kept as $N_A = 10^{14}\,cm^{-3}$, $N_C/N_V = 2 \times 10^{18}\,cm^{-3}$.

10 μm to 300 μm. The maximum value of η is found 9.83% for 100 μm thick commercial available n-cSi wafer. In order to study the effect of metal contact, we have simulated our device taking the front contact of Au metal [16,17]. The thickness of Au contact was kept 70 nm and the n and k values are taken as (1.2152, 1.8434) and (0.155, 3.6024) at 480 nm and 650 nm wavelength, respectively [19,49]. In both wavelengths, the efficiencies were found to be 0.1281%. When the thickness of Au contact was reduced to 2 nm, η reached up to ~ 8% for both wavelengths. This increase in η confirms that 2 nm Au contacts absorb less photon and allow the photons transmission through it to reach the junction. But the decrease in η from 15.51% (Ti contact) to 0.1218% and ~8% (Au contact) might be due to absorption of more photons by Au contact than Ti contact and prevents the photons to pass through it to reach the junction. The further simulation was done by changing the front contact to scandium (n = 1.1198, k = 1.1548 at ~190 nm wavelength [49]. First, the thickness of this contact was considered 2 nm and achieved ~11% efficiency. As this thickness was increased to 10 nm the η was found to decrease up to 0.7633%. We have also optimized the cell parameters considering zigzag phosphorene layer and maximum efficiency of 14.06% was achieved (see Supplementary Figs. S2 and S3).

4. Spectral response, quantum efficiency and temperature dependence

In a solar cell, spectral response (SR) is the ratio of current generated by the solar cell to the optical incident power on the cell and quantum efficiency measures the effects of optical losses such as transmission, reflection etc. through the cell. In order to investigate the spectral response (SR) and quantum efficiencies, simulation was carried out at illumination condition of the monochromatic intensity of $1 \times 10^{16}\,cm^2\,s^{-1}$ and 10 nm spectral widths in the range of 300–1200 nm as mentioned in software. The spectral response (SR), internal quantum efficiency (IQE) and external quantum efficiency (EQE) of the best optimized simulated solar cell (η = 15.51%) are shown in the Fig. 4(a–c). As shown in Fig. 4(a) SR was found to increase from 0.206 to 0.362 A/W as wavelength increased from 300 to 600 nm. Above 600 nm SR decreased and dropped to 0 A/W at 1200 nm. As revealed from magnified Fig. 4(e), photons are absorbed more in 300–600 nm wavelength region which increased J_{SC} in accordance with equation $SR = \frac{J_{SC}(\lambda)}{I_{light}(\lambda)}$, here J_{SC} (λ) is short circuit current density; I_{light} (λ) is intensity of the incident light [50]. In a solar cell, IQE refers to efficiency of photons neither reflected nor transmitted out of the cell, whereas EQE includes the effect of optical losses such as reflection and transmission from the cell. In Fig. 4(b and c) IQE and EQE of the best-optimized cell were found to be 0.852 and 0.851 at 300 nm wavelength. The maximum values of IQE and EQE were found to be 0.915 and 0.914, respectively at 400 nm. Above 400 nm both IEQ and EQE were found to be decreased and reached to 0 at 1200 nm. The IQE is defined as $IQE(\lambda) = \frac{1}{1 - R(\lambda)}\frac{hc}{e\lambda}SR$, [50] which can also be written as $IQE(\lambda) = \frac{1}{A(\lambda) + T(\lambda)}\frac{hc}{e\lambda}SR$, where R (λ) is the reflectivity, h = plank's constant, c = velocity of light, e = electric charge λ = wavelength of incident light (nm) and following the relation used in the software R (λ) + A (λ) + T(λ) = 1, A(λ) = absorptivity and

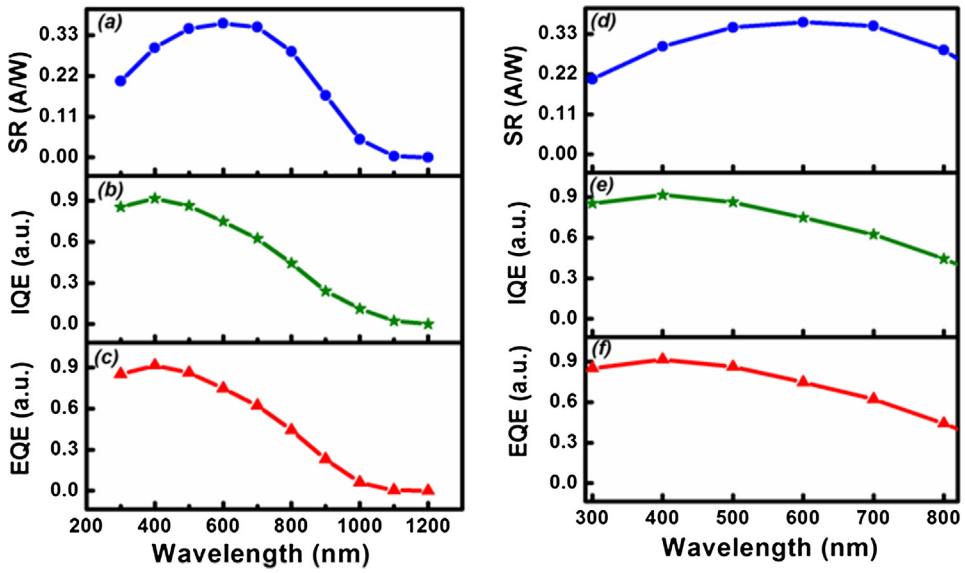


Fig. 4. (a) Spectral response (SR) (b) Internal quantum efficiency (IQE) (c) external quantum efficiency (EQE) within 300–1200 nm range and (d) SR (e) IQE (f) EQE within 300–800 nm range of the best optimized cell.

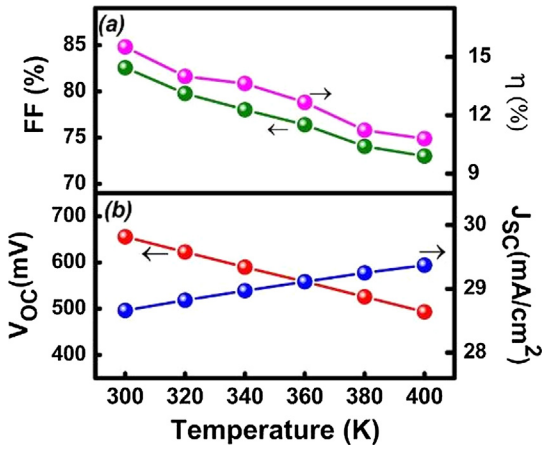


Fig. 5. Variation in cell output parameters η, FF, short circuit current J_{SC} and V_{OC} with temperature change for best optimize cell.

T(λ) = Transmittance. Following IQE vs A(λ), T (λ) and SR, the increase of IQE within the range of 300–400 nm is directly proportional to the SR. But above 400 nm the absorption of photons by the phosphorene layer might be increased and resulted in a decrease of IEQ. Similarly, EQE can be defined as $EQE = \frac{J(\lambda)}{qF_0(\lambda)[1-R(\lambda)]} = \frac{J(\lambda)}{qF_0(\lambda)[A(\lambda)+T(\lambda)]}$, here J (λ) is total current density, $F_0(\lambda)$ is the incident photon density per wavelength interval, q is the total electric charge(q = ne) [51]. From this equation, it can be estimated that in the wavelength range of 300–400 nm, phosphorene layer might absorb less number of photons and allow them to transmit through it to reach the Si layer, which results in an increase of J as well as EQE. After above 400 nm, the phosphorene layer might absorb a large number of photons and prevents them to reach Si layer. As a result, J (λ) as well as EQE decrease. SR, IEQ, and EQE of best optimized zigzag phosphorene/n-cSi (η = 14.06%) solar cell are shown in supplementary Fig. S4(A).

In order to study, the temperature dependence of cell response simulation of the best-optimized cell (η = 15.51%) was carried out in the temperature range from 300 to 400 K, and the results have been shown in Fig. 5(a and b). V_{OC} was found to decrease from 655.5 mV to 493 mV whereas J_{SC} was found to increase slightly from 28.66 mA/cm^2 to 29.37 mA/cm^2 as temperature increases. Since band gap, energy and temperature are related to $E_g(T) = E_g(0) - \left(\frac{\alpha T^2}{\alpha+\beta}\right)$, here E_g (T) is the band gap of the semiconductor at any temperature T, E_g (0) is the band gap energy at T = 0 Kelvin, α and β are the material constant [41]. According to E-T relation as mentioned above, as temperature increases

Table 3
Summary of the parameters of the best optimized cell.

Cell parameters	Phosphorene layer optimized cell (Arm Chair Direction)	Phosphorene layer and n-cSi wafer optimized cell	n-cSi wafer optimized cell for practically available silicon parameters	Experimentally and theoretically available phosphorene based solar cell
V_{OC} (mV)	525.8	655.5	502.3	–
J_{SC} (mA/cm^2)	26.34	28.66	26.06	–
FF (%)	72.48	82.57	75.17	–
η (%)	10.04 (@ Single layer phosphorene at $N_A = 10^{14}$ cm^{-3} $N_c/N_v = 2 \times 10^{18}$ cm^{-3}, $E_g = 1.30$ eV, $\chi = 3.6$ eV, $\varepsilon_r = 2.88$, $\mu_n = 1100$ cm^2/Vs, $\mu_p = 700$ cm^2/Vs, and @ 220 μm n-cSi at $N_D = 9 \times 10^{14}$ cm^{-3} $N_c/N_v = 1 \times 10^{19}$ cm^{-3}, $\chi = 4.22$ eV)	15.51 (@ Single layer phosphorene at $N_A = 10^{14}$ cm^{-3} $N_c/N_v = 2 \times 10^{18}$ cm^{-3}, $E_g = 1.30$ eV, $\chi = 3.6$ eV, $\varepsilon_r = 2.88$, $\mu_n = 1100$ cm^2/Vs, $\mu_p = 700$ cm^2/Vs, and @ 100 μm n-cSi at $N_D = 10^{17}$ cm^{-3} $N_c/N_v = 10^{19}$ cm^{-3}, $\chi = 4.22$ eV)	9.83 (@ Single layer phosphorene at $N_A = 10^{13}$ cm^{-3} $N_c/N_v = 2 \times 10^{18}$ cm^{-3}, $E_g = 1.30$ eV, $\chi = 3.6$ eV, $\varepsilon_r = 2.88$, $\mu_n = 1100$ cm^2/Vs, $\mu_p = 700$ cm^2/Vs, and @ 100 μm n-cSi at $N_D = 10^{15}$ cm^{-3} $N_c/N_v = 2.8 \times 10^{19}/1.04 \times 10^{19}$ cm^{-3}, $\chi = 4.05$ eV and $\mu_p = 450$ cm^2/Vs	η was reported to be in range of 0.3–6.52% (experimentally) [2,52] And 16–18% (theoretically) [53]

band gap decreases. It is also reported that J_o increased as temperature increased [41] and resulted in decrease of V_{OC}. In addition, at lower E_g large numbers of absorbed photons have enough energy to create e-h pairs and J_{SC} gets increased. The FF and η are found to decrease due to decrease of V_{OC}. The maximum efficiency η = 15.51% is found at temperature 300 K. Optimized cell parameters have been summarized in Table 3. The variation of parameters with temperature of the best optimized zigzag phosphorene and n-cSi is shown in supplementary Fig. S4(B).

5. Conclusions

In order to develop a cost-effective heterojunction Si solar cell, the structure Ti/p-phosphorene/n-cSi/Ag has been proposed and simulated using AFROS-HET-v2.5 software to optimize the layer's parameters. Effect of eight parameters of phosphorene was analyzed individually. After optimizing parameters of phosphorene (arm-chair) and n-cSi wafer maximum efficiency η = 15.51% was achieved whereas η = 14.06% in case of zigzag. Our simulation reveals that the efficiency does not depend much on anisotropic nature of phosphorene. Simulated results demonstrate that if the stability of phosphorene in the air can be controlled then it can act as an efficient transparent electrode in the silicon heterojunction solar cell.

Acknowledgment

The authors gratefully acknowledged Helmholtz-Zentrum Berlin for providing AFORS-HET simulation software.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.commatsci.2018.04.059.

References

[1] J.S. Kim, Y. Liu, W. Zhu, et al., Toward air-stable multilayer phosphorene thin-films and transistors, Sci. Rep. 5 (2015) 8989.
[2] S. Fukuoka, T. Taen, T. Osada, Electronic structure and the properties of phosphorene and few-layer black phosphorus, J. Phys. Soc. Jpn. 84 (2015) 121004.
[3] M. Batmunkh, M. Bat-Erdene, J.G. Shapter, Phosphorene and phosphorene-based materials – prospects for future applications, Adv. Mater. 28 (2016) 8586–8617.
[4] J. Qiao, X. Kong, Z.X. Hu, F. Yang, et al., High-mobility transport anisotropy and linear dichroism in few-layer black phosphorus, Nat. Commun. 5 (2014) 4475.
[5] M. Shanmugam, C.A. Durcan, B. Yu, Layered semiconductor molybdenum disulfide nanomembrane based Schottky-barrier solar cells, Nanoscale 4 (2014) 7399–7405.
[6] S. Ge, C. Li, Z. Zhang, Dynamical evolution of anisotropic response in black phosphorus under ultrafast photoexcitation, Nano Lett. 15 (2015) 4650–4656.
[7] Y. Du, L. Yang, H. Zhou, P.D. Ye, Performance enhancement of black phosphorus field-effect transistors by chemical doping, IEEE electron devices letters 37 (2016) 429–432.
[8] Y. He, F. Xa, Z. Shao, J. Zhao, J. Jie, Surface charge transfer doping of monolayer phosphorene via molecular adsorption, J. Phys. Chem. Lett. 6 (2015) 4701–4710.
[9] A. Castellanos-Gomez, L. Vicarelli, E. Prada, et al., Isolation and characterization of few layer black phosphorous, 2D Materials 1 (2014) 025001.
[10] Y. Cai, G. Zhang, Y.W. Zhang, Layer-dependent band alignment and work function of few-layer phosphorene, Sci. Rep. 4 (2014) 6677.
[11] P. Kumar, B.S. Bhadoria, S. Kumar, S. Bhowmick, et al., Thickness and electric-field-dependent polarizability and dielectric constant in phosphorene, Phys. Rev. B 93 (2016) 195428.
[12] W. Hu, L. Lin, C. Yang, J. Dai, et al., Edge-modified phosphorene nanoflake heterojunctions as highly efficient solar cells, Nano Lett. 16 (2016) 1675–1682.
[13] S. Lin, S. Liu, Z. Yang, et al., Solution-Processable Ultrathin Black Phosphorus as an Effective Electron Transport Layer in Organic Photovoltaics, Adv. Funct. Mater. 26 (2016) 864–871.
[14] D.W. Kang, J.Y. Kwon, J. Shim, et al., Al2O3 antireflection layer between glass and transparent conducting oxide for enhanced light trapping in microcrystalline silicon thin film solar cells, Sol. Energy Mater. Sol. Cells 101 (2012) 22–25.
[15] M. Buscema, D.J. Groenendijk, G.A. Steele, et al., Photovoltaic effect in few-layer black phosphorus PN junctions defined by local electrostatic gating, Nat. Commun. 5 (2014) 4651.
[16] J.D. Wood, S.A. Wells, D. Jariwala, et al., Effective passivation of exfoliated black phosphorus transistors against ambient degradation, Nano Lett. 14 (2014) 6964–6970.
[17] E. Tea, C. Hin, Charge carrier transport and lifetimes in n-type and p-type phosphorene as 2D device active materials: *an ab initio study*, Phys. Chem. Chem. Phys. 18 (2016) 22706–22711.
[18] AFORS-HET v2.5, AFORS-HET online help (In-built), Optical Models.
[19] A.D. Rakić, A.B. Djurišić, J.M. Elazar, et al., Optical properties of metallic films for vertical-cavity optoelectronic devices, Appl. Opt. 37 (1998) 5271–5283.
[20] Reflectance Calculator < https://www.filmetrics.com/reflectance-calculator > (accessed February 2017).
[21] Introduction to Quantum Transport in Low-Dimensional Systems. < http://www.hri.res.in/~nanotr16/notes/KJThomas-1.pdf > (accessed March 2017).
[22] Y. Saito, Y. Iwasa, Ambipolar insulator-to-metal transition in black phosphorus by ionic-liquid gating, ACS Nano 9 (2015) 3192–3198.
[23] N. Mao, J. Tang, L. Xie, et al., Optical anisotropy of black phosphorus in the visible regime, J. Am. Chem. Soc. 138 (2016) 300–305.
[24] J. Singh, Semiconductor devices: basic principles, second ed., Wiley India (P) Ltd, New Delhi, 2004.
[25] Electron and Hole Velocities and Mobilities. < http://studyelectronics.blogspot.in/2011/12/electron-and-hole-velocities-and.html > (accessed February 2017).
[26] H. Liu, Y. Du, Y. Deng, Semiconducting black phosphorus: synthesis, transport properties and electronic applications, Chem. Soc. Rev. 44 (2014) 2732–2743.
[27] S. Zhang, J. Yang, R. Xu, et al., Extraordinary photoluminescence and strong temperature/angle-dependent raman responses in few-layer phosphorene, ACS Nano 8 (2014) 9590–9596.
[28] X.Y. Zhou, R. Zhang, J.P. Sun, et al., Landau levels and magneto-transport property of monolayer phosphorene, Sci. Rep. 5 (2015) 12295.
[29] P. Bhaskar, A.W. Achtstein, M.J.W. Vermeulen, et al., Radiatively dominated charge carrier recombination in black phosphorus, J. Phys. Chem. C 120 (2016) 13836–13842.
[30] X. Tian, L. Liu, Y. Du, Variable electronic properties of lateral phosphorene–-graphene heterostructures, Phys. Chem. Chem. Phys. 17 (2016) 31685–31692.
[31] M.T. Edmonds, A. Tadich, A. Carvalho, et al., Creating a stable oxide at the surface of black phosphorus, ACS Appl. Mater. Interface 7 (2015) 14557–14562.
[32] U. Gangopadhyay, S. Roy, S. Garain, et al., Comparative simulation study between n- type and p- type silicon solar cells and the variation of efficiency of n- type solar cell by the application of passivation layer with different thickness using AFORS HET and PC1D, IOSR J. Eng. 2 (2012) 41–48.
[33] Q. Wang, M.R. Page, E. Iwaniczko, Y.Q. Xu, L. Roybal, R. Bauer, B. To, H.C. Yuan, A. Duda, Y.F. Yan, Crystal Silicon Heterojunction solar cells by Hot-Wire CVD, Presented at the 33rd IEEE Photovoltaic Specialists Conference, San Diego, California, May 2008.
[34] A. Datta, P. Chetterjee, Computer Modeling of Heterojunction with Intrinsic Thin Layer "HIT" Solar cells: Sensitivity Issues And Insights Gained, in: Prof. Leonid A. Kosyachenko (Ed.), Solar Cells - Thin-Film Technologies, InTech, 2011, pp. 275–302 (Chapter 13).
[35] A.V. Sachenko, V.P. Kostylyov, M.V. Gerasymenko, et al., Analysis of the silicon solar cells efficiency. Type of doping and level optimization, Semiconduct. Phys., Quant. Electron. Optoelectron. 19 (2016) 67–74.
[36] S. Zhong, X. Hua, W. Shen, Simulation of high-efficiency crystalline silicon solar cells with homo-hetero junctions, IEEE Trans. Electron Dev. 60 (2013) 2104–2110.
[37] X. Miao, S. Tongay, M.K. Petterson, et al., High efficiency graphene solar cells by chemical doping, Nano Lett. 12 (2012) 2745–2750.
[38] S. Xu, X. Zeng, W. Wang, et al., Simulation and optimization characteristic of novel MoS2/c-Si HIT solar cell, J. Minerals Mater. Charact. Eng. 5 (2017) 323–338.
[39] Notes on Heterostructure Fundamental. < https://engineering.purdue.edu/~ee606/downloads/T3.PDF > (accessed February 2017).
[40] G.J.H.M. Wilfried, V. Sark, et al., Physics and technology of amorphous crystalline heterostructure silicon, Solar Cells (1997) 197–395.
[41] K.K. Nair, J. Jose, A. Ravindran, Analysis of temperature dependent parameters on solar cell efficiency using MATLAB, IJEDR 4 (2016) 536–541.
[42] M. Belarbi, M. Beghdad, A. Mekemeche, Simulation and optimization of n-type interdigitated back contact silicon heterojunction (IBC-SiHJ) solar cell structure using Silvaco Tcad Atlas, Sol. Energy 127 (2016) 206–215.
[43] Y. Zhao, C. Liang, M. Sun, et al., Effect of doping on the short-circuit current and open-circuit voltage of polymer solar cells, J. Appl. Phys. 116 (2014) 154506.
[44] Y. Zhou, G. Long, The low density of conduction and valence band states contribute to the high open voltage in perovskite solar cells, J. Phys. Chem. C 121 (2017) 1455–1462.
[45] H. Rashid, K.S. Rahman, M.I. Hossain, et al., Prospects of Molybdenum Disulfide (MoS2) as an alternative absorber layer material in thin film solar cells from numerical modeling, Chalcogenide Lett. 11 (2014) 397–403.
[46] M.Y. Khairy, A.G.N. Salwa, S.L. Cheow, A.W. Shahrul, M.Y. Othman, M. Y. Sulaiman, A. Zaharim, K. Sopian, Enhancing Efficiency of Thin Film Silicon Solar Cell Published in the Proceedings of the 3rd WSEAS Int. Conf. on Renewable Energy Sources, Greece, 2009.
[47] C.J.J. Tool, A.R. Burgers, P. Manshanden, A.W. Weeber, Effect of wafer thickness on the performance of mc-si solar cells, Presented at 17th European Photovoltaic Solar Energy Conference, Munich, 2001.
[48] S.M. Sze, Physics of Semiconductor Devices, second ed., A Willey Interscience Publication, Singapore, 1981.
[49] Refractive index, INFO, < https://refractiveindex.info/?shelf=main&book=Ni&page=Johnson > (accessed September 2017).
[50] W.J. Yang, Z.Q. Ma, X. Tang, et al., Internal quantum efficiency for solar cells, Sol. Energy 82 (2008) 106–110.
[51] H. Bashiri, M.A. Karami, S.M. Nejad, Crystalline silicon solar cell engineering to improve fill factor, open circuit voltage, short circuit current and overall cell efficiency, Modares J. Electr. Eng. 14 (2016) 6–14.
[52] Y. Deng, Z. Luo, N.J. Conrad, et al., Black phosphorusmonolayer MoS2 van der Waals heterojunction p-n diode, ACS Nano 8 (2014) 8292–8299.
[53] J. Dai, X.C. Zeng, Bilayer Phosphorene: effect of stacking order on bandgap and its potential applications in thin-film solar cells, J. Phys. Chem. Lett. 5 (2014) 1289–1293.

Manvendra Verma et al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 5, Issue 1, March 2018, pp. 201-205

Geopolymer concrete: A way of sustainable construction

Manvendra Verma[1], & Dr Nirendra Dev[2]

[1]Department of Civil Engineering, Delhi Technological University, Delhi, India.
[2] Professor & Head, Department of Civil Engineering, Delhi Technological University, Delhi, India.

Abstract: **Geopolymer concrete is an eco-friendly, economical, cementless, and durable concrete. By the comprehensive survey of the literature; Geopolymer concrete is sustainable development for the construction industry. Geopolymer concrete reduces the carbon footprints by the using industrial solid waste material like as Flyash and blast furnace slag as a binding material. Around one ton of carbon dioxide (CO_2) emission for the production of the one ton of cement. In geopolymer concrete, we use the waste material like as flyash and slag as a binding material. Moreover, that binding material activating by the alkaline solution (sodium hydroxide and sodium silicate). Around 120 million tonnes of Flyash and 12 million tonnes of slag produced annually at the thermal power station and Steel Plant respectively in India. Presently in India, due to Limited modes of practices of utilisation use amount of slag and flyash dumped in lands of each production unit and engaging in valuable agricultural land and grave pollution to the whole environment. Hence, Geopolymer concrete is a perfect alternative to Portland cement concrete.**

Keywords: **Fly ash, Slag, Geopolymer concrete, Alkaline activator, Portland cement.**

I. INTRODUCTION

Concrete is the world's most reliable, durable, versatile construction material. After to water, Concrete is the user most material which required large quantities of Portland cement. Portland cement production is the second to the automobile as the dominant generator of Co_2, which polluted production. So, required to find an alternative material to the existing most expensive most resource consuming Portland cement. Geopolymer concrete is an innovative material which is produced by the chemical action of inorganic molecules. Geopolymer concrete a way to the sustainable development of concrete Industries. Geopolymer concrete made by the industrial solid waste like as flyash or slag, an alkaline solution, aggregate, sand, water and superplasticiser. Industrial solid waste like as slag and flyash used as a binding material and alkaline solution used for the activating the industrial solid waste for geopolymerization. The industrial solid waste contains silica and Alumina; geopolymer concrete reduces the carbon footprints due to reducing the use of Portland cement. In 1 ton production of cement produces the one ton of CO_2 in the atmosphere. Around 7 to 8% of CO_2 produced by the cement production in the world. CO_2 is a Greenhouse gas that increases the temperature of the earth. In the development of countries increase the development of the infrastructure that is built by the concrete. So, production of concrete increases day by day due to increase in the demand for infrastructure. Around 120 million tons of flyash produces from the thermal power plant and 12 million tons of slag generated from a steel plant in India. Many valuable cultivated land used for the disposal of the slag and flyash. Industrial solid waste polluted the soil and the underground of that area. By the use of industrial solid waste, we did the double profit work one by the use of waste to prevent the land from the disposal or dumping and second by the generating the concrete.

II. GEOPOLYMER CONCRETE

Geopolymer concrete is a new kind of concrete which uses different chemistry to that found in Portland cement concrete. Davidovits (1991) firstly introduce the word 'Geopolymer' to sketch an alternative cementitious material which has ceramic-like properties. As opposed to OPC, the manufacture of geopolymer does not consume high levels of energy. A geopolymer concrete is made of the amorphous alumino-Silicate materials such as flyash and slag with alkali based Chemicals such as sodium silicate and sodium hydroxide. Geopolymer concrete does not need to OPC to work. Geopolymer is inorganic polymers that have been phased attracting world attention to potentially Revolution material. Similar to nature zeolite minerals, geopolymer is a class of 3 dimensionally networked alumino-silicate materials. A wide range of natural aluminium- silica minerals such as flyash and slag could serve as potential source material for the synthesis of geopolymers.

Geopolymer is a type of amorphous aluminosilicate cementitious material that exhibits the ideal property of Rock-forming minerals due to hardness, chemical stability and longevity. Geopolymerization involves a heterogeneous chemical reaction between solid(binder powder) and Aluminium silicate oxide and alkali metals silicate solution at highly alkaline conditions and mild temperature yielding three-dimensional amorphous to semicrystalline polymeric and ring structure which consists of Si-O-Al and Si-O-Si bonds. the systematic formation of polycondensation material by alkali into poly(sialate-siloxo) can be shown as described by the equation

$n(Si_2O_5 , Al_2O_2) + nH_2O + nSi_2$
(NaOH, KOH)

$n\ (OH)_3 – Si\ -O-Al(OH)_2 – O-\ Si(OH)_3 -$
(NaOH, KOH)

© 2018 IJRAA All Rights Reserved

(Na^+, K^+)-(Si-O-Al-O-Si-O-) + nH_2O

III. LITERATURE REVIEW

Patimapon Sukman et al. examine the short-term resistant ability of two different cementitious system prepared using silty clay as a significant component of five wt.% sodium sulphate and 5wt.% magnesium sulphate solution. The cementitious system is clay Portland cement and clay high calcium flyash geopolymer concrete. Moreover, concluded that the exposure to magnesium sulphate causes more degradation than sodium sulphate in both clays flyash geopolymer and clay cement system. The 28 days strength of clay fly ash geopolymer is up to 1.2 times higher than the clay cement.

Weibo Ren et al. Study the residual behaviours of alkali-activated flyash slag geopolymer concrete under different heating temperature and cooling regimes. Analysis the change in weight, compressive strength, UPV test, wavelet test, and microstructure of concrete. Compressive strength, UPV of FSDC decrease due to increase in temperature especially under extreme heat condition. The specimen cooled by watering is most degradable than the specimen cooled naturally.

Mohammad Amer Rafique Bhutta et al. Examine the performance of geopolymer concrete prepared using pulverised fuel(PFA) and palm oil fuel ash from Agro-Industrial waste along with alkaline Activator. Analysis the mass changes, compressive strength, mineralogical property, microstructural changes by exposed to 5% magnesium sulphate solution for up to 18 months. The resulted that OPC concrete shows mass decrease up to 20% compared to 4% decreased for blended fuel ash geopolymer concrete. The Portland cement concrete was more susceptible to attack by 5% sodium sulphate solution compared to Blended fuel ash geopolymer concrete. The performance of geopolymer concrete in 5% sodium sulphate solution was more stable due to cross-linked aluminosilicate polymer structure as compared to normal portland cement C-S-H structure.

William Gustavo et al. Examine the performance of geopolymer concrete based on a binary mixture of flyash and ground granulated blast furnace slag(GGBFS) in 80/20 ratio and activated with a mixture of sodium silicate and sodium hydroxide. Analysis the durability of FA/GGBFS and OPC concrete by immersed in 5 wt % sodium sulphate and magnesium sulphate solution. Check the volumetric expansion, mineralogical characteristics, microstructural changes and compressive strength of concrete. Moreover, found that the loss of compressive strength in FA/GGBFS system was 33% for 360 days of magnesium sulphate exposure and 48% of the Portland cement concrete. The expansion of FA/GGBFS concrete and OPC concrete were .041 and 0.086 respectively for magnesium sulphate exposure. The expansion of FA/GGBFS concrete and OPC concrete is 0.0068% and 0.412% respectively in sodium sulphate solution.

Farhad Aslani investigated the relation for normal and high strength geopolymer concrete at elevated temperature. Moreover, to establish the efficient Modelling and specify the fire performance criteria of concrete structures. Analyse the compressive strength, flexure strength, thermal strain and modulus of elasticity of concrete. He proposed the free thermal strain relationship for unstressed and prestressed geopolymer concrete at a higher temperature. He proposed the relationship between the compressive strength, elastic modulus and flexure strength of unstressed geopolymer concrete with different geopolymer resins composition geopolymeric Binder type and aggregate type at elevated temperature. Moreover, the results are in good reasonable agreement with the experimental results.

Paulo H.R. Bourges et al. Studied that to improve the geopolymer formulation by the development of micro concrete. He investigated the three parameters one) the Andreasen packing factor q= .21, .235, and .26 which are alter particle size distribution of aggregate, 2) solution to solid rate 1.3,1.4, and 1.5, 3) type of aggregate either glass or quartz. Moreover, concluded that geopolymer concrete made with Quartz present better properties than made with glass aggregate. To make self-compacting mix uses the intermediate Andreasen factor q= 2.235 and solution to the solid ratio are 1.4.

DV Reddy et al. Evaluates the durability characteristics of low calcium fly ash based geopolymer concrete that was subjected to a corrosive Marine environment. Test the geopolymer concrete beams which contained flyash with 8M and 14M concentration of sodium hydroxide and sodium silicate solution. Moreover, conclude that the strength increases between 7 to 28 days are 15% for 8M concentration and 7% for 14M concentration geopolymer concrete and 33% increases of OPC concrete. The splitting tensile strength was also consistently higher for geopolymer concrete comparative to OPC concrete. The electrical resistivity of geopolymer concrete was not significantly affected, and reduced cracking implies a permeability reduction.

Partha Sarathi Deb et al. Studied the effect of ultrafine flyash on setting time, strength and porosity of geopolymer paste and specimen cured at 20^0C made with 15% GGBFS or 10% OPC. Analyse setting time, compressive strength, mineralogical properties, and microstructure of concrete. They improve the density of microstructure by using ultra-fine flyash that reduced the porosity and increase compressive strength. They found that the maximum compressive strength at 10% ultra-fine flyash mix with flyash only geopolymer and 5% ultrafine flyash for GGBFS or OPC blended geopolymer. Using of ultrafine flyash reduce d setting time of flyash geopolymer paste.

Chaicharn Chotetanorm et al. Studied that resistance to sulphate attack, compressive strength, sorptivity, and pore size of high calcium bottom ash in geopolymer mortar. They used the ground lignite bottom ash with medium particle sizes of 16, 25 and 32 micrometres in a mortar. They use sodium hydroxide and sodium silicate and temperature curing to activate the geopolymerization. Moreover, concluded that the fine bottom Ash gives the high compressive strength, low rate of sorptivity and low expansion in a mortar, When subjected to 10% sodium sulphate solution for 360 days.

Pattanopong Topark Ngarm et al. Investigated the strength, setting time, and bond of high calcium fly ash based geopolymer concrete. They use Sodium Hydroxide with 10M, 15M, and 20M concentration, sodium Syndicate to Sodium Hydroxide ratio of 1, 2, alkali liquid flyash ratio of 0.5 and

© 2018 IJRAA All Rights Reserved

two curing regimes heat curing at 60+2^{0}C for 24 hours and room temperature curing at 23+2^{0}C.were used. Moreover, concluding that high-calcium flyash suitable for producing high strength geopolymer concrete with high Bond strength between concrete and rebar. The optimum sodium oxide content was approximately 12% by weight of flyash, and optimum strength of 54.4MPa was obtained with the 15M sodium hydroxide. Bond strength of high calcium geopolymer concrete was slightly higher than low calcium fly ash based geopolymer concrete.

ST Erdogan studied the strength development of mixture containing only ground perlite as their powder(Binder) activated with sodium hydroxide(NaOH) and sodium silicate solution at room temperature or using oven dry curing. XRD, FTIR, and NMR analysis indicate that when cured at 100^{0}C sodium hydroxide solution activated mixture undergoes more geopolymerization then we activated mixture. Activator gave high strength mixture 40-50MPa when samples cured in a dry oven at 100^{0}C when both activator have used the contribution of WC to strength development is Greater.

P. Chindraprasirt et al. Studied the synthesis of high strength geopolymer using fine high calcium fly ash. Analyse the effect of fineness of flash on the setting time of geopolymer paste workability, strength development and dry shrinkage geopolymer mortar made from classified fine high calcium fly ash. To find the optimum point of delay time, Curing temperature and duration of curing at high temperature. Moreover, found that the optimum parameter for fine flyash activated with sodium silicate solution and sodium hydroxide ten molar solution at 1:1 mass ratio was 1 hour for delay time and 3-day heat curing at 75^{0}C. The high calcium fly ash based geopolymer mortar continue to gain strength when kept in a normal atmospheric condition after the initial heating curing period.

Gaurav Nagalia et al. Investigated the role of alkali hydroxide and its concentration on the microstructure and compressive strength of fly ash based geopolymer concrete. Geopolymer concrete with different alkaline solutions[NaOH, KOH, $Ba(OH)_2$ LiOH] was prepared by mixing Class C (9.42% CaO), and class f fly ash (1.29% CaO). Moreover, concluded that the higher concentration of CaO in flyash provide higher strength in the finished geopolymer concrete. Sodium Hydroxide was the only one that resulted in mechanical strength that is comparable to Portland cement concrete.

Kunal Kupwade Patil et al. Investigated the alkali-silica reaction between reactive aggregates and geopolymer Matrix. The specimen was prepared by using Class C and class f flyash stockpiles. OPC concrete exhibited higher average expansion by a factor of six compared to geopolymer concrete specimens. alkali-silica reaction product was formed on Portland cement concrete specimen on the geopolymer concrete specimen made with Class C flyash. Fly ash-based geopolymer concrete is significantly less vulnerable to ASR than Portland cement based concrete. Elevated temperature results in reinitiation of the geopolymerization process of unreacted flyash particles leading to lower porosity and higher strength.

Kunal Kupwade Patil et al. Investigated the durability of Steel reinforced concrete specimen made from the alkali-activated flyash stockpiles and OPC in cyclic wet-dry chloride environment was evaluated over a period of 12 months. Analyse the chloride diffusion, corrosion potential and rates Pore structure and chloride content, microstructure and chemical analysis of the specimen. Moreover, concluded that geopolymer concrete specimen exhibit lower average diffusion, coefficients, chloride content and porosity compared with their Portland cement counterparts. Geopolymer concrete specimen made from class f flyash exhibited significantly higher resistance to chloride-induced corrosion compared to OPC as well as class f fly ash concrete.

Biruk Hailu Tekle et al. Investigated the bond property between sand coated GFRP bars a corrosive resistant substituent to a Steel bar and flyash based geopolymer concrete. They analyse the bond failure mode, bond relationship curve, effect of embedment length, the effect of compressive strength and bar diameter. Moreover, concluded that the bond performance of Sand coated GFRP reinforced geopolymer concrete is better than the Portland cement concrete. The pull out load increases with increasing amendment length, but the average Bond strength decreases. The residual Bond stress is caused by the friction resistance between the concrete and sand coating of GFRP bars. Load increases the bond stress in the vicinity of free end increases because of the redistribution of bond stress along the amendment length.

M. Albitar et al. Investigated the behaviour of reinforced concrete at all load level and understanding of the bond between the reinforcement and concrete is required. The bond between the concrete and the reinforcement control the formation of cracks widening and stiffening at the serviceability limit. Analyse the pullout test, accelerated corrosion method, sorptivity and water absorption test, frictional strength, maximum slip and corrosion induced cracking. Moreover, concluded that corrosion of reinforcement not only reduce the strength but also lead to deterioration of the bond which can cause increased deflection, reduced strength and ultimate failure. A new model of geopolymer concrete with new corroded enforcement was proposed.

O. Sanusi et al. Investigated the effect of replacing some percent(%) of natural coarse aggregate with recycled aggregate in geopolymer concrete synthesis. In this study focus on the compressive strength and release of elements from the geopolymer concrete sample. Moreover, concluded that use of recycled aggregate in geopolymer concrete affect both the compressive strength and leaching of the elements. The compressive strength was reduced for recycled aggregate addition with rates UP to approximately 30% but then slightly increased for more magnificent addition. Mixes with less than 30% recycled coarse aggregate released at a level higher than the screening limit. Incremental addition of recycled coarse aggregate leads to a reduction of as leaching. TCLP testing indicated that most of the COPC were not leached at a level above the EPA soil screening levels or the TCLP regulatory threshold.

Ginghis Maranan et al. Investigator job on the performance of sand coated GFRP bars of a parameter such as a bar diameter 12.7mm, 15.9mm and 19 mm and the embedment length (5, 10 and 15 diameters of bars). The test of the direct

© 2018 IJRAA All Rights Reserved

pullout test of GFRP bars into geopolymer concrete with a compressive strength of 33Mpa. Moreover, concluded that the specimen with a longer embedment length failed because of concrete splitting whereas the specimen with a shorter embedment length failed because of the bar pull out. Specimen failed because of bar diameter cause the peak average bond stress to decrease. The increase in the embedment length results in lower average Bond stress because of the different type failure.

Zhen Liu et al. Investigated the problems caused by the construction and demolition waste and the depletion of natural aggregate in the present. Both the course recycled aggregate and the fine recycled aggregate are used to produce new green concrete with fly ash based geopolymer concrete. moreover, concluded that the compressive strength, young modulus, and poisons ratio decrease with the increase of water content and w/c ratio for the geopolymer recycled aggregate concrete specimens. young modulus is more vulnerable to the change of w/c ratio than compressive strength. SEM and Nanoindentation Technologies so there is no interfacial transition between the old cement paste. The fly ash based geopolymer is capable rendering a lower w/c ratio for the production of recycled aggregate concrete than Portland cement while even producing a higher compressive strength.

Cheah Chee Ban et al. Investigated a new geopolymer system based on hybridisation between high calcium wood ash and pulverised fuel as which uses as a low % of Alkaline activator (5% of binder wt.), ambient temperature curing and pressurised forming a method for the fabrication of load-bearing masonry unit. the removal of alkali Hydroxide was made possible by the arcanite minerals (approximately 12%) which are inherently present in HCWA. moreover, concluded that the early strength development is due to primarily by the formation of N-A-S-H and K-A-S-H geopolymeric gel. Meanwhile, the long-term strength development is caused by the formation of (Ca, K)-A-S-H geopolymeric gel and secondary C-S-H gels.

IV. CONCLUSION

Based on the discussion, it is concluded that geopolymer concrete has excellent potential to use as a construction material in several applications. Geopolymer concrete is a perfect alternative to Portland cement concrete. In a comprehensive survey of papers, we found that the geopolymer concrete has excellent physical and mechanical properties as compared to Portland cement concrete. Geopolymer concrete durability is higher as compared to the Portland cement. Geopolymer concrete highly resistive to sulphate attack, acid attack, alkali-silica reaction. Geopolymer concrete indirectly reduces the carbon footprints by the usage of industrial solid wastes and reduce the usage of portland cement. Because in the production of portland cement high amount of carbon dioxide generated. So, geopolymer concrete is a way of sustainable developments.

V. REFERENCES

[1]. Chaicharn Chotetanorm, Prinya Chindaprasirt, Vanchai Sata, Sumrerng Rukzon and Apha Sathonsaowaphak, “High-Calcium Bottom Ash Geopolymer: Sorptivity, Pore Size, and Resistance to Sodium Sulfate Attack” J. Mater. Civ. Eng., 2013, 25(1): 105-111.

[2]. P. Chindaprasirt, T. Chareerat, S. Hatanaka and T. Cao, “High-Strength Geopolymer Using Fine High-Calcium Fly Ash” J. Mater. Civ. Eng., 2011, 23(3): 264-270.

[3]. Kunal Kupwade-Patil and Erez N. Allouche, “Impact of Alkali-Silica Reaction on Fly Ash-Based Geopolymer Concrete” J. Mater. Civ. Eng., 2013, 25(1): 131-139.

[4]. O. Sanusi, B. Tempest, V. O. Ogunro and J. Gergely, “Leaching Characteristics of Geopolymer Cement Concrete Containing Recycled Concrete Aggregates” J. Hazard. Toxic Radioact. Waste, 2016, 20(3): -1—1.

[5]. Pattanapong Topark-Ngarm, Prinya Chindaprasirt and Vanchai Sata, “Setting Time, Strength, and Bond of High-Calcium Fly Ash Geopolymer Concrete” J. Mater. Civ. Eng., 2015, 27(7): -1—1.

[6]. Weibo Ren, Jinyu Xu and Erlei Bai, “ Strength and Ultrasonic Characteristics of Alkali-Activated Fly Ash-Slag Geopolymer Concrete after Exposure to Elevated Temperatures” J. Mater. Civ. Eng., 2016, 28(2): -1—1.

[7]. Patimapon Sukmak, Pre De Silva, Suksun Horpibulsuk and Prinya Chindaprasirt, “Sulfate Resistance of Clay-Portland Cement and Clay High-Calcium Fly Ash Geopolymer” J. Mater. Civ. Eng., 2015, 27(5): -1—1.

[8]. Muhammad Aamer Rafique Bhutta, Warid M. Hussin, Mohd Azreen and Mahmood Mohd Tahir, “Sulphate Resistance of Geopolymer Concrete Prepared from Blended Waste Fuel Ash” J. Mater. Civ. Eng., 2014, 26(11): -1—1.

[9]. Farhad Aslani, “Thermal Performance Modeling of Geopolymer Concrete” J. Mater. Civ. Eng., 2016, 28(1): -1—1.

[10]. William Gustavo Valencia Saavedra, Daniela Eugenia Angulo and Ruby Mejía de Gutiérrez, “Fly Ash Slag Geopolymer Concrete: Resistance to Sodium and Magnesium Sulfate Attack” J. Mater. Civ. Eng., 2016, 28(12): -1—1.

[11]. Zhen Liu, C. S. Cai, Hui Peng and Fenghong Fan, “Experimental Study of the Geopolymeric Recycled Aggregate Concrete” J. Mater. Civ. Eng., 2016, 28(9): -1—1.

[12]. Kunal Kupwade-Patil and Erez N. Allouche, “Examination of Chloride-Induced Corrosion in Reinforced Geopolymer Concretes” J. Mater. Civ. Eng., 2013, 25(10): 1465-1476.

[13]. Partha Sarathi Deb and Prabir Kumar Sarker, “Effects of Ultrafine Fly Ash on Setting, Strength, and Porosity of Geopolymers Cured at Room Temperature” J. Mater. Civ. Eng., -1—1.

[14]. Cheah Chee Ban, Part Wei Ken, and Mahyuddin Ramli, “Effect of Sodium Silicate and Curing Regime on Properties of Load Bearing Geopolymer Mortar Block” J. Mater. Civ. Eng., -1—1.

[15]. D. V. Reddy, Jean-Baptiste Edouard, and Khaled Sobhan, “Durability of Fly Ash–Based Geopolymer

© 2018 IJRAA All Rights Reserved

Structural Concrete in the Marine Environment" J. Mater. Civ. Eng., 2013, 25(6): 781-787.

[16]. Gaurav Nagalia, Yeonho Park, Ali Abolmaali and Pranesh Aswath, "Compressive Strength and Microstructural Properties of Fly Ash-Based Geopolymer Concrete" J. Mater. Civ. Eng., 2016, 28(12): -1—1.

[17]. Paulo H. R. Borges, Lucas F. Fonseca, Vitor A. Nunes, Tulio H. Panzera, and Carolina C. Martuscelli. "Andreasen Particle Packing Method on the Development of Geopolymer Concrete for Civil Engineering" J. Mater. Civ. Eng., 2014, 26(4): 692-697.

[18]. S. T. Erdogan "Properties of Ground Perlite Geopolymer Mortars" J. Mater. Civ. Eng., 2015, 27(7): -1—1.

[19]. Biruk Hailu Tekle, Amar Khennane, and Obada Kayali. "Bond Properties of Sand-Coated GFRP Bars with Fly Ash-Based Geopolymer Concrete" J. Compos. Constr., 2016, 20(5): -1—1.

[20]. M. Albitar, P. Visintin, M. S. Mohamed Ali, O. Lavigne, and E. Gamboa (2016). "Bond Slip Models for Uncorroded and Corroded Steel Reinforcement in Class-F Fly Ash Geopolymer Concrete" J. Mater. Civ. Eng., -1—1.

[21]. Ginghis Maranan, Allan Manalo, Karu Karunasena, and Brahim Benmokrane. "Bond Stress-Slip Behavior: Case of GFRP Bars in Geopolymer Concrete" J. Mater. Civ. Eng., 2015, 27(1): -1—1.

[22]. Manvendra Verma, Nirendra dev (2017). "Review on the effect of the different parameter on the behaviour of Geopolymer concrete." IJIRSET.2017.0606210.

IJRRA

© 2018 IJRAA All Rights Reserved

Indian J Microbiol
https://doi.org/10.1007/s12088-018-0744-y

SHORT COMMUNICATIONS

GroEL Mediates Folding of *Bacillus anthracis* Serine/Threonine Protein Kinase, PrkC

Richa Virmani[1,2] · Yogendra Singh[2] · Yasha Hasija[1]

Received: 16 May 2018 / Accepted: 23 May 2018
© Association of Microbiologists of India 2018

Abstract *Bacillus anthracis* causes anthrax in human and animals. Both, signaling system such as two component system and endogenous chaperone system such as GroEL–GroES help bacteria to cope with the environmental challenges. Such molecular chaperones are the stress induced proteins that help bacteria to override unfavorable conditions by their moonlighting functions. Previous reports showed that PrkC and PrpC, the Ser/Thr kinase–phosphatase pair in *B. anthracis*, control phosphorylation of GroEL and regulate biofilm formation. In this study, we show that GroEL is involved in the folding of PrkC to active form. The proteins (GroEL, PrkC and PrpC) were expressed and purified by affinity chromatography. Purified GroEL was used for refolding of denatured PrkC and PrpC and observed that GroEL refolds PrkC but not PrpC as measured by their enzymatic activity. We also observed that purification of GroEL with six histidine tag using Cobalt-Agarose resin yielded superior quality GroEL protein with negligible contamination of non-specific proteins. Thus, cobalt resin can be a better choice for purification of many histidine tagged proteins, where Ni-NTA does not work very well.

Keywords *Bacillus anthracis* · GroEL · PrkC · Signaling · Protein folding

✉ Yogendra Singh
ysinghdu@gmail.com

✉ Yasha Hasija
yashahasija@gmail.com

[1] Delhi Technological University, New Delhi, Delhi 110042, India

[2] Department of Zoology, University of Delhi, New Delhi, Delhi 110007, India

Bacillus anthracis is the causative agent of anthrax which makes spores in stress conditions. The disease is initiated when spores enter in the host, germinate in macrophages and start producing anthrax toxin. During the life cycle (vegetative cell and spore), *Bacillus* comes across various environmental stresses where heat shock proteins (hsps) play an important role in their survival [1]. Hsps comprise a set of molecular chaperones that transiently assist in repairing of misfolded and partially unfolded proteins. Folding of proteins is an error prone mechanism that requires a thermodynamic balance to drive the protein towards its native state [2]. The GroEL–GroES chaperonin system is one of the most well studied examples of molecular chaperone in bacteria which is also conserved in *B. anthracis* [3].

Homo-oligomeric GroEL forms two heptameric rings stacked back to back forming a cylindrical structure that takes the specific substrate proteins into a central cavity [4]. The structure of GroEL is divided into three subunits with apical, intermediate and equatorial domains. The exposed hydrophobic residues on the misfolded or unfolded substrate proteins are captured by the apical domains of GroEL thus enclosing the substrate into a cage like structure. A ring like cofactor, GroES, then covers this central cavity. When the folding process is complete, the folded protein and GroES gets dissociated from GroEL [5]. GroEL is one such immunogenic Hsp, which is located on the exo-sporium and upon immunization provides protection against *B. anthracis* infection [6]. GroEL in combination with protective antigen (PA) enhanced protective efficacy of PA against anthrax spore challenge [7].

B. anthracis possess an intricate intracellular signaling machinery to survive in extreme environmental stress conditions [8–11]. The kinase phosphatase pair, PrkC and PrpC are the conserved mediators of bacterial signaling

Published online: 30 May 2018

system which regulate critical cellular processes like sporulation, germination, cell growth and virulence in *B. anthracis* [8, 9]. Previous studies have shown that PrkC mediated phosphorylation of GroEL is associated with biofilm formation in *B. anthracis* [3]. GroEL has specificity for its substrates as it assists in the folding of nearly 10% of the newly synthesized polypeptides. Substrate proteins interact with GroEL not only during the initial folding process but also for maintenance of conformation upon exposure to different environmental conditions during their lifetime [12]. Here, we studied whether PrkC and PrpC is a substrate of GroEL–GroES system in *B. anthracis*. We purified GroEL from *B. anthracis* and analyzed its role in folding of heat denatured PrkC and PrpC proteins by measuring their enzymatic activity.

We used an already available expression vector pYS5 which harbors a native promoter of PA that expresses the downstream gene constitutively [13]. To study the protein folding potential of native GroEL purified from *B. anthracis*, the gene for GroEL was cloned downstream of PA promoter with a His tag using primers listed in Table 1. The lysate of *B. anthracis* Sterne cells (Bas) and Bas transformed with pYS5-*groel*-His_6 plasmid (Bas His GroEL) were resolved on SDS-PAGE and subjected to immunoblotting using anti-GroEL antibodies (α-GroEL antibody; 1: 15,000 dilution and HRP-conjugated anti-mouse IgG antibody; Bangalore Genei, 1: 18,000 dilution) to confirm overexpression (Fig. 1a). Ni^{2+} and Co^{2+} immobilized on a matrix was employed to purify native GroEL. Ni^{2+}-NTA matrix has a high binding affinity for the hexa-histidine tag at pH 8.0 and Co^{2+} resin display weaker yet specific interaction for hexa-histidine tag [14]. The lysate prepared from *B. anthracis* Sterne cells harboring pYS5-*groel*-His_6 was passed through the matrix followed by washing and elution with a buffer containing imidazole. For purification from Co^{2+} matrix, the cell pellet was lysed in a buffer [1X Tris-buffered saline (TBS), 1 mM PMSF, 1 mg/mL lysozyme, and PI (Roche)] through sonication. The lysate collected after harvesting was incubated with Co^{2+} superflow resin for 3 h. After washing with wash buffer (1X TBS, 1 mM PMSF, and 10 mM imidazole) for 30 min, the protein was eluted in elution buffer (1X TBS, 1 mM PMSF, and 300 mM imidazole). To check the purity of the protein, the eluted fractions from both resins (Ni^{2+}-NTA and Co^{2+} matrix) were resolved on SDS-PAGE and immunoblot was probed with anti-His antibody (Abcam 1: 18,000 dilution). The blot was developed by Super Signal® West Pico Chemiluminescent Substrate kit (Pierce Protein Research Products) according to the manufacturer's protocol [15].The results showed that GroEL was purified to homogeneity using Co^{2+} matrix while the protein purified from Ni^{2+} matrix showed a few contaminating bands (Fig. 1b).

Our previous work has identified GroEL as a substrate of PrkC, which plays an important role in bacterial biofilm formation [3]. It is known that GroEL acts on unfolded or incorrectly folded proteins to fold them in correct conformation. PrkC is known to be a stable enzyme at relatively higher temperatures, which potentiates its role in bacterial spores. Hence, the refolding capability of GroEL can be measured on the basis of activity of its substrates after refolding. The sequences encoding *prkC* catalytic domain (*bas3713*, 1-337aa), *groes* (*bas0252*) were amplified using gene specific primers (Table 1) containing appropriate restriction sites from *B. anthracis* Sterne strain genomic DNA [16–19]. The amplified products were then digested and ligated into an *E. coli* expression vector pProEx-HTc digested with the same enzymes. The His tagged proteins

Table 1 Primers and clones

Genes	Vectors	Restriction sites	Primers (5′ → 3′)[a]	References
prkC (*bas3713*, 1-1011 bp), kinase domain	pProEx-HTc	FP-BamHI	TAGGTGAAGTGGATCCTGCTGATTGGAAAACGC	[3]
		RP-XhoI	TGTAATTAAAATCTCGAGTCATTTATTACTTCGTTTG	
prpC (*bas3714*)	pProEx-HTc	FP-BamHI	GCGCAAAGAAGAGACGAGGGATCCAGATGAAAGCCGTGTTTCT	[3]
		RP-XhoI	CGTTTTCCAATCAGCACGCTCGAGTTCACCTACTTTCGTTTGTCG	
groES (*bas0252*)	pProEx-HTc	FP-BamHI	AACAAAATGAGGAGGATCCTG TTCATGCTAAAG	[3]
		RP-XhoI	CCTTGGATTTCTCGAGATTTATATAATTAACCG	
groEL-His	pYS5	FP-SpeI	CCAAGGGGGTCAACTAGTTATGGCAAAAGATATTAA	This study
		RP-BamHI	GGATCCTTAGTGGTGATGGTGGTGGTGCATCATTCCGCCCATAC	
groEL (*bas0253*)	pProEx-HTc	FP-BamHI	AATCCAAGGGGGTGGATCCTTATGGCAAAAG	[3]
		RP-XhoI	TTAGGGCAAACTCGAGTTACATCATTCCGCCC	

[a]Restriction/mutation sites have been underlined

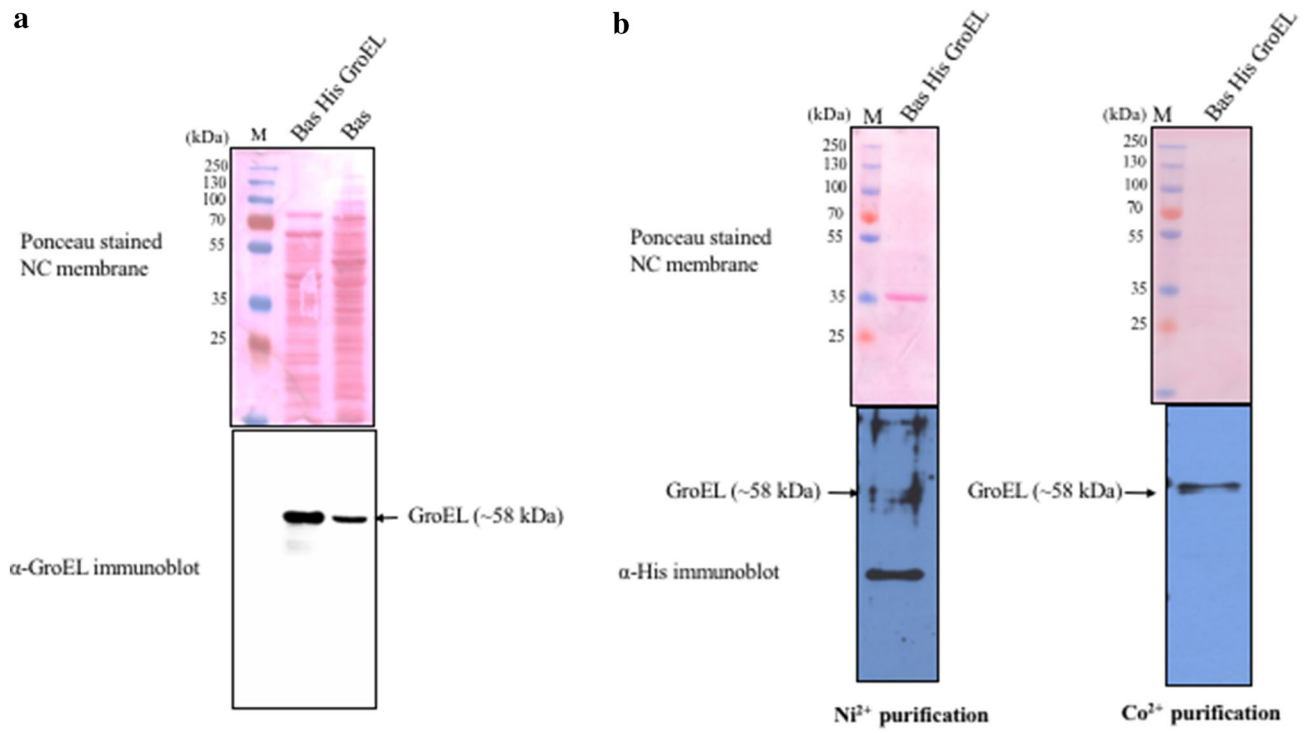


Fig. 1 Purification of GroEL using $Ni^{2+}$and Co^{2+} resin. **a** *B. anthracis* Sterne lysate expressing GroEL was resolved on SDS-PAGE and transferred to nitrocellulose membrane. Upper panel depicts Ponceau S stained membrane to show total protein loaded in the well and lower panel shows image of immunoblot using anti-GroEL antibodies. **b** The Bas His GroEL purified from Ni^{2+} and Co^{2+} resins were resolved on SDS-PAGE followed by immunoblotting using anti-His antibodies

(PrkC and GroES) were over-expressed in *E. coli* BL-21 (DE3) [18–20] and affinity purified with Ni^{2+}-NTA resin (Qiagen). PrkC (0.5 μM) was denatured by heat treatment and incubated with Bas His_6-GroEL (2 μM) and GroES (1 μM) at room temperature. Enzymatic activity of PrkC was quantitated by evaluating its autophosphorylation in a kinase assay [20]. Purified PrkC was taken as a positive control and denatured PrkC (incubated with GroES alone) was taken as negative control for measuring the relative activity. The relative phosphorylation of denatured PrkC was compared considering autophosphorylation of purified PrkC as 100% (Fig. 2a). The results show that the denatured PrkC upon incubation with GroEL gets refolded and regained its activity to about 60% as compared to the purified PrkC (Fig. 2b). This result suggests that GroEL refolds PrkC as measured by kinase activity.

Our previous work also identified that *B. anthracis* phosphatase PrpC dephosphorylates GroEL which decreases its oligomerization [3]. PrpC was also analyzed as a substrate of GroEL. The gene, *prpC* (*bas3714*) was amplified using gene specific primers and purified using Ni^{2+}-NTA resin. As with PrkC, PrpC (1.8 μM) was also heat denatured and then incubated with Bas His_6-GroEL (2 μM) and GroES (1 μM) for refolding. The refolding of PrpC protein was assessed by measuring its phosphatase activity using pNPP as a substrate. The reaction was carried out in a 96-well plate format and the plate was incubated at 37 °C for 30 min followed by measurement of absorbance at 405 nm (Microplate reader, Bio-Rad). Native PrpC was taken as positive control and GroEL–GroES mixture was taken as negative control. The results show that GroEL was unable to refold PrpC (Fig. 3).

These results suggest that GroEL may help PrkC to remain in an active confirmation during the transition stages such as sporulation and infection. We purified GroEL with Cobalt beads and found that PrkC lost > 85% of activity after denaturation, which was substantially regained after incubation with native GroEL. On the other hand, PrpC did not refold in the presence of GroEL, indicating that it is not a GroEL substrate. Thus, role of GroEL in regulating signal transduction proteins involved in bacterial pathogenesis indicates utilization of chaperones by microbial pathogens in their virulence. Further, characterization of the signaling proteins may help to explore the potential role of GroEL as a drug target.

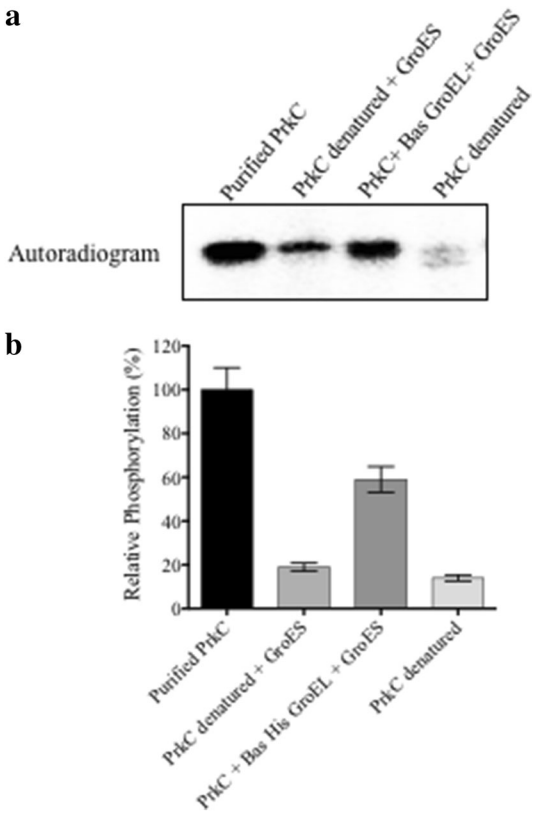


Fig. 2 GroEL mediated refolding of PrkC regains its activity. Denatured PrkC was allowed to refold in the presence of Bas His GroEL and its activity was assessed by in vitro kinase assays. Autophosphorylation of purified PrkC was taken as 100% and relative phosphorylation of denatured PrkC was calculated. After denaturation, PrkC retained only ~ 10% activity, which was regained to 60% in the presence of GroEL and GroES. All the refolding assays were performed in the presence of GroES. Upper panel **a** shows the autoradiogram whereas, lower panel **b** shows the densitometric analysis

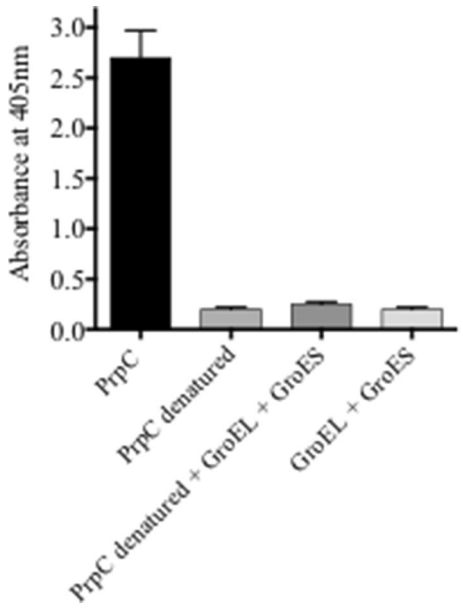


Fig. 3 Refolding of PrpC by GroEL: Denatured PrpC was incubated with Bas His GroEL in the presence of GroES and activity of PrpC was measured at 37 °C using pNPP as a substrate. Activity of purified PrpC was taken as positive control and samples without PrpC was taken as negative control

Acknowledgements We thank Dr. Gunjan Arora for his valuable suggestions.

Authors' Contributions RV conceived and designed the experiments. YH and YS contributed materials, reagents and analysis tools. All authors have read and approved the manuscript.

Funding CSIR senior research fellowship to R.V. is acknowledged.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Neckers L, Tatu U (2008) Molecular chaperones in pathogen virulence: emerging new targets for therapy. Cell Host Microbe 4:519–527. https://doi.org/10.1016/j.chom.2008.10.011
2. Guerriero CJ, Brodsky JL (2012) The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. Physiol Rev 92:537–576. https://doi.org/10.1152/physrev.00027.2011
3. Arora G, Sajid A, Virmani R, Singhal A, Kumar CS, Dhasmana N, Khanna T, Maji A, Misra R, Molle V, Becher D (2017) Ser/Thr protein kinase PrkC-mediated regulation of GroEL is critical for biofilm formation in *Bacillus anthracis*. NPJ Biofilms Microbiomes 3:7. https://doi.org/10.1038/s41522-017-0015-4
4. Van Duijn E, Heck AJ, Van Der Vies SM (2007) Inter-ring communication allows the GroEL chaperonin complex to distinguish between different substrates. Protein Sci 16:956–965. https://doi.org/10.1110/ps.062713607
5. Madan D, Lin Z, Rye HS (2008) Triggering protein folding within the GroEL–GroES complex. J Biol Chem 283:32003–32013. https://doi.org/10.1074/jbc.M802898200
6. Redmond C, Baillie LW, Hibbs S, Moir AJ, Moir A (2004) Identification of proteins in the exosporium of *Bacillus anthracis*. Microbiology 150:355–363. https://doi.org/10.1099/mic.0.26681-0
7. Sinha K, Bhatnagar R (2013) Recombinant GroEL enhances protective antigen-mediated protection against *Bacillus anthracis* spore challenge. Med Microbiol Immunol 202:153–165. https://doi.org/10.1007/s00430-012-0280-z
8. Shah IM, Laaberki MH, Popham DL, Dworkin J (2008) A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. Cell 135:486–496. https://doi.org/10.1016/j.cell.2008.08.039
9. Bryant-Hudson KM, Shakir SM, Ballard JD (2011) Autoregulatory characteristics of a *Bacillus anthracis* serine/threonine kinase. J Bacteriol 193:1833–1842. https://doi.org/10.1128/JB.01401-10
10. Arora G, Sajid A, Arulanandh MD, Singhal A, Mattoo AR, Pomerantsev AP, Leppla SH, Maiti S, Singh Y (2012) Unveiling the novel dual specificity protein kinases in *Bacillus anthracis* identification of the first prokaryotic dual specificity tyrosine phosphorylation-regulated kinase (DYRK)-like kinase. J Biol

Chem 287:26749–26763. https://doi.org/10.1074/jbc.M112.351304
11. Sajid A, Arora G, Singhal A, Kalia VC, Singh Y (2015) Protein phosphatases of pathogenic bacteria: role in physiology and virulence. Annu Rev Microbiol 69:527–547. https://doi.org/10.1146/annurev-micro-020415-111342
12. Ye X, Mayne L, Kan ZY, Englander SW (2018) Folding of maltose binding protein outside of and in GroEL. Proc Natl Acad Sci 115:519–524. https://doi.org/10.1073/pnas.1716168115
13. Singh Y, Chaudhary VK, Leppla SH (1989) A deleted variant of *Bacillus anthracis* protective antigen is non-toxic and blocks anthrax toxin action *in vivo*. J Biol Chem 264:19103–19107
14. Jiang C, Wechuck JB, Goins WF, Krisky DM, Wolfe D, Ataai MM, Glorioso JC (2004) Immobilized cobalt affinity chromatography provides a novel, efficient method for herpes simplex virus type 1 gene vector purification. J Virol 78:8994–9006. https://doi.org/10.1128/JVI.78.17.8994-9006.2004
15. Sajid A, Arora G, Gupta M, Singhal A, Chakraborty K, Nandicoori VK, Singh Y (2011) Interaction of *Mycobacterium tuberculosis* elongation factor Tu with GTP is regulated by phosphorylation. J Bacteriol 193:5347–5358. https://doi.org/10.1128/JB.05469-11
16. Jamal MA, Sharma SP, Chung HJ, Kim HJ, Hong ST, Lee S (2017) Ultra-high efficient colony PCR for high throughput screening of bacterial genes. Indian J Microbiol 57:365–369. https://doi.org/10.1007/s12088-017-0665-1
17. Gupta S, Tewatia P, Misri J, Singh R (2017) Molecular modeling of cloned *Bacillus subtilis* keratinase and its insinuation in psoriasis treatment using docking studies. Indian J Microbiol 57:485–491. https://doi.org/10.1007/s12088-017-0677-x
18. Arora G, Sajid A, Singhal A, Joshi J, Virmani R, Gupta M, Verma N, Maji A, Misra R, Baronian G, Pandey AK (2014) Identification of Ser/Thr kinase and forkhead associated domains in *Mycobacterium ulcerans*: characterization of novel association between protein kinase Q and MupFHA. PLoS Negl Trop Dis 8:e3315. https://doi.org/10.1371/journal.pntd.0003315
19. Singhal A, Arora G, Virmani R, Kundu P, Khanna T, Sajid A, Misra R, Joshi J, Yadav V, Samanta S, Saini N (2015) Systematic analysis of mycobacterial acylation reveals first example of acylation-mediated regulation of enzyme activity of a bacterial phosphatase. J Biol Chem 290:26218–26234. https://doi.org/10.1074/jbc.M115.687269
20. Arora G, Sajid A, Arulanandh MD, Misra R, Singhal A, Kumar S, Singh LK, Mattoo AR, Raj R, Maiti S, Basu-Modak S (2013) Zinc regulates the activity of kinase-phosphatase pair (BasPrkC/BasPrpC) in *Bacillus anthracis*. Biometals 26:715–730. https://doi.org/10.1007/s10534-013-9646-y

2018 CoNECD - The Collaborative Network for Engineering and Computing
Diversity Conference: Crystal City, Virginia Apr 29

#ILookLikeAnEngineer: Using Social Media Based Hashtag Activism Campaigns as a Lens to Better Understand Engineering Diversity Issues

Dr. Aqdas Malik, George Mason University
Dr. Aditya Johri, George Mason University

Aditya Johri is Associate Professor in the department of Information Sciences & Technology. Dr. Johri studies the use of information and communication technologies (ICT) for learning and knowledge sharing, with a focus on cognition in informal environments. He also examine the role of ICT in supporting distributed work among globally dispersed workers and in furthering social development in emerging economies. He received the U.S. National Science Foundation's Early Career Award in 2009. He is co-editor of the Cambridge Handbook of Engineering Education Research (CHEER) published by Cambridge University Press, New York, NY. Dr. Johri earned his Ph.D. in Learning Sciences and Technology Design at Stanford University and a B.Eng. in Mechanical Engineering at Delhi College of Engineering.

Mr. RAJAT HANDA
Mr. Habib Karbasian, George Mason University

PhD student in IT

Dr. Hemant Purohit, George Mason University

c American Society for Engineering Education, 2018

#ILookLikeAnEngineer: Using Social Media Based Hashtag Activism Campaigns as a Lens to Better Understand Engineering Diversity Issues

Abstract

Each year, significant investment of time and resources is made to improve diversity within engineering across a range of federal and state agencies, private/not-for-profit organizations, and foundations. In spite of decades of investments, efforts have not yielded desired returns - participation by minorities continues to lag at a time when STEM workforce requirements are increasing. In recent years a new stream of data has emerged - online social networks, including Twitter, Facebook, and Instagram - that act as a key sensor of social behavior and attitudes of the public. Almost 87% of the American population now participates in some form of social media activity. Consequently, social networking sites have become powerful indicators of social action and social media data has shown significant promise for studying many issues including public health communication, political campaign, humanitarian crisis, and, activism. We argue that social media data can likewise be leveraged to better understand and improve engineering diversity. As a case study to illustrate the viability of the approach, we present findings from a campaign, #ILookLikeAnEngineer (using Twitter data – 19,354 original tweets and 29,529 retweets), aimed at increasing gender diversity in the engineering workplace. The campaign provided a continuous momentum to the overall effort to increase diversity and novel ways of connecting with relevant audience. Our analysis demonstrates that diversity initiatives related to STEM attract voices from various entities including individuals, large corporations, media outlets, and community interest groups.

Introduction

The term "STEM education" refers to teaching and learning in the fields of science, technology, engineering, and mathematics. According to Sanders (2008), in the 1990s, the National Science Foundation (NSF) "SMET" was the shorthand for "science, mathematics, engineering, and technology" and an NSF program officer complained that "SMET" sounded too much like "smut," that resulted in the new acronym "STEM" [1]. Although it took some time for STEM to catch on, and even as recently as 2003 few people know what it meant, in the last decade the use of the term has exploded [2]. STEM is now used to refer to educational activities across all grade levels (from pre-school to post-doctorate) and to both formal (e.g., classrooms) and informal (e.g., after school programs) experiences. According to data available publicly, annual federal appropriations for STEM education are now in above $4 billion. There is an active interest in STEM education at that level given the importance placed on advancing science and engineering to maintain U.S. competitive edge. STEM education is now a topic used frequently in policy debates, including issues of education, workforce development, national security and immigration. According to the Congressional Research Service STEM primer [3], there are between 105 and 252 STEM education programs or activities at 13 to 15 federal agencies; the key agencies involved in the effort are Department of Education, National Science Foundation, and Health and Human Services.

Given that the interests are broad and federal efforts are spread across multiple agencies, there is a concern with perceived duplication of effort and a lack of coordination in the federal effort. Therefore, efforts to improve accountability and coordination have gained prominence in recent years. The data currently available about STEM education paints a "complicated" picture [3]. According to many indicators [3], overall graduate enrollments in science and engineering (S&E) grew 35% over the last decade; S&E enrollments for Hispanic/Latino, American Indian/Alaska Native, and African American students (all of whom are generally underrepresented in S&E) grew by 65%, 55%, and 50%, respectively. Yet, concerns remain about persistent academic achievement gaps between various demographic groups, STEM teacher quality, the rankings of U.S. students on international STEM assessments, foreign student enrollments and increased education attainment in other countries, and the ability of the U.S. STEM education system to meet domestic demand for STEM labor.

To better understanding STEM efforts, different kinds of data sources are used, such as National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), Program for International Student Assessment (PISA), Science and Engineering Indicators (S&E), and NCSES, among others. GAO and other federal offices also periodically release information that assists in understanding STEM issues. The use of surveys is the preferred method to produce data about different programs and initiatives both due to the convenience of data collection and because the data collection and analysis is consistent with statistical measures that have been perfected over decades. The data and analysis generated here is reliable and useful. But, as Feuer (2013) has argued, in spite of this data and the effort to better understand STEM outcomes "the connections between diagnosis and prescription were, and remain, logically and empirically challenging (p. 3)." **Traditional data are limited in what they tell us about STEM education.** In particular, they do not capture the pulse of the people nor provide insight into the major entities that can assist STEM initiatives. They are, like most STEM initiatives, top-down and very strongly tied to activities in formal institutions. Consequently, there has been a renewed

interest in recent years to revisit the limitations and quality of existing data and, as Feuer (2013) argues, "the need for more and better data, and priority topics for a sustained research agenda (p. 5)" [4].

These newer data and information are needed especially if we want to pursue and achieve many of the goals outlined by policy researchers that are shown to work. For instance, Malcom-Piqueux & Malcom (2013) argue that greater awareness of careers, role models, after- and out-of-school experiences are critical for equity. They suggest that "students and their families need encouragement and access to information at a much earlier stage than has typically been provided, through exposure to role models who look like them, information about the kinds of jobs done by persons with preparation in engineering, and examples to dispel the idea that engineering is solitary work (pg. 32) [5]." They further contend that efforts to increase the participation of females and underrepresented minorities in engineering education and careers started outside of school, through programs such as Expanding Your Horizons for female students and Mathematics, Engineering, Science Achievement (MESA), Sally Ride Science and AAAS' Spark Club (part of a suite of programs under NSF's Innovative Technology Experiences for Students and Teachers, ITEST), After School Alliance, and others but few parents have awareness of these programs and communities must be engaged and relevant contexts created for students from different populations. All the issues outlined here are easy to address if we have the reach – access to students and their families – and if we are able to provide them with the right information at the right time. Furthermore, the efforts we make will be even more fruitful if we are able to construct channels that are reusable, through which we can easily access the right target population, and which are natural for our intended audience.

In recent years, with the advent of the Internet, new sources of data about public behavior and discourse of public interactions are increasingly becoming available. This increase in the quantity and quality of data has occurred at the same time as the capacity to collect and store data as well the ability to analyze this data have improved significantly. One significant source of data for understanding events and activities that are of concern and of issues that matter to us is data generated from social networking sites or social media. Social media include Internet-based platforms such as Facebook, Instagram, and Twitter, among others, which continue to gain users due to increase in mobile devices based user interaction through different apps that are available as part of the device ecosystem. Almost 87% of the American population now participates in some form of social media activity. Social media provide powerful indicators of social action as they allow individual and small groups to find like-minded individuals and organizations and topics around which they can coalesce. Social media data has been especially useful for researchers studying issues around health communication [6], politics [7], humanitarian crisis [8], social movements such as Arab Spring and Black Lives Matters [9], [10] and activism [11]. For example, Achrekar et al. (2011) has exploited Twitter data for a real-time assessment to predict flu. The model is trained and tested with the data collected by U.S. Center for Disease Control and Prevention, and shows that Twitter data can improve the model prediction accuracy [12]. These efforts suggest that social media is an appropriate channel for better understanding STEM education issues.

STEM education has a key implication for the workforce development and thus, the distribution of STEM workforce can be a reflection of the state of STEM education in the society.

Unfortunately, the recent events and campaigns for stereotyping in the workforce indicate the poor diversity of STEM workers and how the minorities are stereotyped – particularly women. This behavior further undermines the efforts to raise awareness for STEM education. On the other hand, there are several initiatives by established organizations to raise awareness for workforce diversity and thus, implicitly promoting STEM education. In this paper we draw on content analysis and social network analysis methods to conduct an exploratory study of STEM workforce stereotyping and diversity issues using social media activities, and participating audience in a popular campaign. We apply data, techniques, expertise, and interdisciplinary perspectives that have not been applied to the STEM domain by exploring a STEM diversity initiate – #ILookLikeAnEngineer campaign. Using a bottoms-up approach and looking at informal influences on STEM education by means of awareness for workforce diversity, including the media, public, and the businesses, we build a more comprehensive and nuanced picture.

The fundamental contribution of this research is in shedding light on issues that have the potential to impact multiple stakeholders and how they pursue STEM related efforts. For instance, for projects looking for STEM partners in a specific geographical area, our research can provide that information. There is an opportunity to aggregate information about scholarships and competitions as they are often announced through social media. Overall this exploratory research can help build a system that will assist in raising awareness of STEM issues, better understanding of issues, better understanding of sentiment regarding issues, figure out geographic distribution, provide temporal understanding of issues, and help understand communities and networks – who is connected and who can be connected. For instance it will provide an avenue to share real world examples [13]; allow participants to share their STEM related stories [14]; and, engender better debates around STEM issues [15].

Hashtag Activism Campaign: #ILookLikeAnEngineer

In August 2015 a hashtag on Twitter, #ILookLikeAnEngineer, started trending highly and attracted a lot of media coverage and participation by users. The hashtag was born out of one woman's frustration with how people reacted to a recruitment advertisement by her company One Login that featured her. In late July 2015 the company OneLogin posted billboards across public transport in the California Bay Area, especially at the BART train stations, as a recruitment drive (see Fig. 1). One of the billboards depicted a woman engineer Isis Wenger (now Isis Anchalee) and her photo attracted a lot of attention on the Web. Her image led to discussions online about the veracity of the campaign as some people found it unlikely that she was really an engineer.



Fig. 1. OneLogin billboard at a BART station

Online comments stated that she was "too attractive" to be a "real engineer," among a lot of other demeaning comments. These online discussions, according to Anchalee, prompted her to write a post on Medium, on August 1, 2015, responding to the stereotypical reaction her image in the ad had generated. She stated in the post that, "At the end of the day, this is just an ad campaign and it

is targeted at engineers. This is not intended to be marketed towards any specific gender — segregated thoughts like that continue to perpetuate sexist thought-patterns in this industry." She wanted to show that she was not the only female engineer and also take some of the attention away from her and towards the issue of diversity in engineering. Her Medium post as well the hashtag received large support and the Twitter tag in particular took off as a way for others to share their images and their ideas on the issue. The hashtag soon saw significant media coverage not just in the US but also across the Atlantic, particularly in the UK. The movement not only had an online presence but a large event was organized in September 2015 in the Bay Area supported by a group. The hashtag movement was started as a campaign to raise funds to be able to post billboards across the Bay Area with the hashtag and photos; in effect, to take the online movement into the physical world through a series of billboards.

Research Study

Analytical Foci, Research Questions, and Overview of Approach

Our exploratory research combines existing research on STEM Education issues, such as broadening participation, with new techniques in the emerging field of computational social science [16] for analysis of data, for visualization of data, and for overall improving how data can be used. Our early-concept has two analytical foci (see Table 1 below for details):

1) *Actors*: Who are the participants in STEM workforce diversity related tweets? For this question, we use a qualitative coding scheme to categorize types of participating actors, and study patterns of their interaction in the network.

2) *Activity*: What activities do they engage in and what is the distribution of these activities? For this question, we use geographical/spatial-temporal analysis to gauge the interest towards STEM workforce diversity efforts and campaigns related messages from specific regions, over time.

Table 1. Overview of Research Study – Analytical Foci, Expected Contribution, and Approach

Analytical Foci and Purpose	**Contribution to Improve Understanding of STEM Issues**
Actors	
Who participates? Identify primary individuals and organizations participating in STEM issues.	Establish who is participating and who is absent – individuals as well as organizations. Understand the nature of participation and self-representation of participants in order to leverage them for interventions.
Activities	
What do participants do? Identify major STEM related activities, initiatives, and opportunities.	It is important that any activity or opportunity is appropriately targeted and there is outreach around it. It is also important to know, geographically and in terms of topics, what is going on when and where.

Research Data

For this research we use data from the social network site Twitter - a highly popular microblogging platform. Twitter, in addition to Facebook, has become the de facto tool for direct sharing of information by members of the public and by organizations. Through the use of hashtags, tweeting, and retweeting, Twitter allows users to voice their concerns, connect with others and show solidarity, while also demonstrating their personal and/or organizational identities.

Our overall data corpus though will be larger than just the number of tweets as we will also collect publicly available information referenced in the tweets. For instance, users often refer to external webpages and images which can provide useful information about the topic. There is also some limited but useful information available about what users and the public at large searches for on the Web that can be usefully linked to data from social media. This part is quite important as in addition to the words in the tweets, the links when expanded lead to more important information. In spite of the advances in data science, there are multiple challenges in working with any form of data, including Twitter data, and we are well aware of some of the issues we might face. For instance, the data from Twitter are often noisy, grammatically incorrect, and short in length, which challenges the direct application of standard text analytics methods. To undertake meaningful analysis of such data is not straightforward. For our analysis we will leverage studies in other domains as many relevant techniques have been developed and we will develop some domain specific algorithms if required.

We initiated the data collection for the dataset through Twitter streaming API, by seeding three hashtags for crawling: #ILookLikeAnEngineer, #LookLikeAnEngineer, and #LookLikeEngineer. The time frame for the data ranged from August 3rd, 2015 (the day hashtag was first used) till October 15th, 2015. In order to have a reliable and stabilized retweet count, we recollected the tweets through Twitter search API after a year. The final dataset consists of 19,354 original tweets and 29,529 retweets. Along with tweet text, the following metadata was also collected: retweet count, favorite count, time of tweet; for the user: name, screen name, location, followers, following, likes, and number of tweets.

Analysis

Even though our intent in this proposed work is to undertake research which is exploratory and risky, to test the viability of our approach we conducted a preliminary analysis. *Exploratory research questions* that guided the preliminary study included: a) Participation: Who are the most active users in terms of follower count in the ILookLikeAnEngineer (ILLAE) dataset? Who are the most active users in terms of tweet count? Where are the tweets coming from? b) Discourse: What are the most frequent hashtags and their associated themes? c) Interaction: What is a pattern of user interaction and its dynamics? In the following sections we present the detailed analysis of these inquiries.

Participation

User analysis:

To understand the drivers of the campaign, we looked at the aspects related to highly significant tweeters with respect to number of tweets, retweets, and favorites count. In the current dataset, 13,270 unique users contributed through their tweets generating a total of 29,529 retweets and

43,450 favorites. In order to gain a better understanding of the prominent users within the network, we specifically focused on the top 25 users. Statistics of the user profile also provided insights that lead to a better understanding of the drivers of the campaign and their intentions. We further classified each of the respective Twitter profile manually by browsing through the Twitter page.

Analysis of the most active users indicates active involvement of female professionals (mostly engineers) and large engineering companies. Seven women were in the top 25 users with highest number of tweets. Michelle Glauser who was the campaign manager of #ILookLikeAnEngineer, followed by a Female IT engineer from India tweeted the most throughout the campaign (see Table 2). The campaign originator Isis Anchalee as well as women engineers from large corporations such as GE and Cisco also contributed actively to the campaign. From the corporate side, GE was one of the leading entities as they actively tweeted using five different Twitter accounts: GEConnections, GECareers, GeneralElectric, GEHiresHeroes, and GEEuropeCareers. Other engineering companies active in the campaign included Caterpillar and Cisco.

Table 2. Users posting highest number of tweets

Twitter handle	**User type**	**No. of tweets**	**No. of Followers**
Michelle Glauser	Female	106	3845
Aditi Kapoor	Female	89	317
Richa Sachdevaa	Female	65	182
Isis Anchalee	Female	56	20017
GEConnections	Company	52	17266
Sitesetup	Company	47	1132
Shaunda	Female	45	710
Caterpillar Inc.	Company	37	96244
Hackathon News	Company	36	2039
Savy Nic	Male	32	338
SNC-Lavalin	Company	31	16524
GECareers	Company	31	22738
General Electric	Company	27	426703
GEHiresHeroes	Company	26	2659
We Are Cisco	Company	24	22540
Brian Ballantyne	Male	24	516
GE Europe Careers	Company	24	7571
Zzadet	Individual	24	104
Suzanne Defache	Female	24	390
Nasser Sahlool	Male	23	596
ThoughtWorks	Company	23	161701
Nevin Weston	Female	23	242
CiscoWISE	Company	21	1128
BookBuddi	Company	20	285
Stockmarketnow	Company	20	971

Even though organizations participated actively in terms of the number of tweets (see Table 2), their presence was limited in other indices of engagement. In spite of a large number of followers only six organizations make up in the list of top 20 retweets (see Table 3). This finding indicates

the limited impact of corporates in engaging users to retweet their content within the campaign. On the other hand, tweets by individuals were retweeted substantially more than other entities. Individuals accounted for 60% of the most retweeted tweets highlighting the influence of individuals in hashtag activism campaigns. Though some of these individuals are well known such as Scott Kelly, Tracy Chou, and Isis Anchalee (campaign founder), yet many of the users are not (e.g. Marcos Caceres and Jolene Hayes).

Further examination of corporations in the list reveals that out of the six corporations identified, five are media outlets. We also analyzed the ratio of verified and non-verified entities to assess their respective level of engagement. The analysis indicates that slightly over half of the accounts were verified (55%), and the rest of them (45%) were non-verified accounts. In general, non-verified users on Twitter have a limited reach in terms of number of followers, yet within this campaign their tweets were retweeted proactively that signifies their importance in propagating the message in such campaigns. One of the reasons behind this phenomenon might be that tweets from non-verified users (who are generally individuals) have an emotional tone in their tweets. Furthermore, it is also plausible that the users who retweeted tweets of non-verified individuals might have associated themselves personally and emotionally with the campaign.

Table 3. Users receiving highest number of retweets

Twitter handle	**User type**	**Verified/Non-verified**	**Tweet type***	**No. of retweets**
Dara Oke	Female	NV	T+P	2450
Marcos Caceres	Male	NV	T+P	2088
Keyz MIning	Male	NV	T+L	1999
Isis Anchalee	Female	NV	T+P	898
TechCrunch	Company	V	T+L	803
The New York Times	Company	V	T+L	771
BuzzFeed	Company	V	T+P+L	715
Tesla	Company	V	T+P	713
MIT	University	V	T+P	710
Tracy Chou	Female	V	T+P	670
Scott Kelly	Male	V	T+P	660
Mashable	Company	V	T+P+L	659
BuzzFeed	Company	V	T+P+L	654
Kaya Thomas	Female	V	T+P	588
Erica Joy	Female	V	T+P	586
Pamela Assogba	Female	NV	T+P	552
SciencePorn	Community	NV	T+P+L	512
Lisa Smith	Female	NV	T+P	494
Holly	Female	NV	T+P	491
Jolene Hayes	Female	NV	T+P	450

*T=text, P=photo, L=Link

Likewise retweets count, the favorites count also shows a comparable result. The list of 20 most favorited tweets indicates that 60% of them were tweeted by individuals (see Table 4). Organization, media, NGO/community, and university hold rest of the 40%. Moreover, tweets of

verified and non-verified accounts are favorited almost equally. Overall our findings from the most favorited and retweeted tweets are in agreement that, for the online discourse related to this campaign individuals (ordinary citizens and to some extent notable personalities) dominated over other entities including organizations and media outlets.

Table 4. Users receiving highest number of favorites

Twitter handle	**User type**	**Verified/Non-verified**	**Tweet type**	**No. of favorites**
Marcos Caceres	Male	NV	T+P	3378
Dara Oke	Female	NV	T+P	2725
Scott Kelly	Male	V	T+P	2011
Tracy Chou	Female	V	T+P	1427
Tesla	Company	V	T+P	1373
Isis Anchalee	Female	NV	T+P	1264
Erica Joy	Female	V	T+P	1264
BuzzFeed	Company	V	T+P+L	964
Jolene Hayes	Female	NV	T+P	954
Lisa Smith	Female	NV	T+P	940
BuzzFeed	Company	V	T+P+L	927
Emily Calandrelli	Female	V	T+P+L	912
Kaya Thomas	Female	V	T+P	900
Snapchat	Company	V	T+P	889
Pamela Assogba	Female	NV	T+P	888
SciencePorn	Community	NV	T+P+L	861
Texas Instruments	Company	NV	T+P	839
Ford Motor Company	Company	V	T+P	803
MIT	University	V	T+P	797
Sailor Mercury	Female	NV	T+P	772

*T=text, P=photo, L=Link

In order to get a better sense of participating entities within the campaign, we further examined the engagement level of various entities. We analyzed the 284 most retweeted user profiles out of 13270 unique user profiles (~2.14%). These user profiles were classified into five categories: individuals (e.g. Isis Anchalee, Tracy Chou, and Erica Joy), companies (e.g. Tesla, Google, and Microsoft), news and media outlets (e.g. CNN, BBC, and Independent), non-government organizations/communities (e.g. Code for America, Care.org, and Women's Engineering Society) and universities (e.g. MIT, UC San Francisco, and Victoria). The results from this analysis further reaffirm our previous finding as an overwhelming majority of individuals compose this list (42.6%) followed by media/news outlets (22.5%), and organizations (19.4%).

Further analysis indicates that among the 284 most retweeted users over 46% (N=131) are non-verified. Despite a general assumption that non-verified users are not public figures or well-known organizations, and have far limited number of followers, they were actively retweeted and favorited. This finding also endorses the notion that this campaign has been driven by individuals who identified with the core idea behind the movement. On the other hand, involvement of verified users was equally significant for the success of the campaign. A large number of corporates and

news/media outlets participating in this campaign have verified accounts indicating their support in wider propagation of the campaign message.

Geolocation analysis:

Analyzing the tweets location further helped us in gauging the engagement level of the users based on their location. A limited number of tweets (0.77%) in the collected dataset had geolocation metadata. Therefore, in order to have a better understanding of the location of users within our dataset, we captured the location of the users through location information their Twitter profiles. Out of the total of 19,354 original tweets, around 25% of the user profiles did not display the location. Overall 2600 unique locations were geocoded through Google Map API. The granularity of the results goes up to the county, state, or city level depending on how detailed the location is provided by the user. By parsing the JSONs, we were able to find the coordinates of the user locations that are displayed in the world map (see Fig. 2).

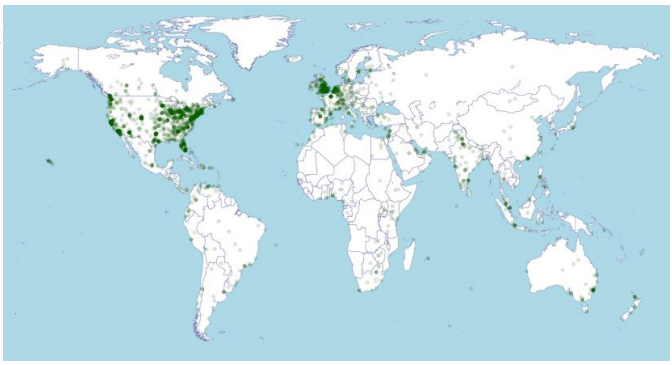


Fig. 2. Tweet activity across the globe

Analyzing the geolocation indicates that users across the globe participated in the campaign (138 countries). An overwhelming majority of the tweets (64%) were from US-based users. Twitter users from Canada, UK, and European countries (predominantly France and Switzerland) with a relatively large engineering and technical presence also actively participated in the campaign. As majority of the tweets were from US-based users, we limited our detailed analysis to state level. The analysis indicate that majority of the tweets were sent from east and west coast, particularly from the states where most of the high-tech companies are based.

Campaign discourse

Words and hashtags analysis

To understand the discussion topics and themes of interest dueing the campaign, we carried out words and hashtags analysis. For words analysis, we utilized Latent Dirichlet Allocation (LDA) which is an unsupervised machine learning technique [17]. This technique has been often used and considered highly suitable for identifying topics related to tweets. We used log likelihood of the model for each topic. We found out that log likelihood for six topics was the highest and was sufficient to see what all is being discussed related to hashtag. As a complement to the words analysis, we also utilzed the term frequency analysis and visulaized it with the top term used within the tweets (see Fig. 3).



Fig. 3. Keywords used in #ILLAE tweets.

Results from LDA technique revealed five key discussion themes. The main topic of interest as captured by keywords: *community, tech, inspiring, hosted,* and *powerful* refer to the event that was hosted on August 13 to raise funds for the campaign billboards to be erected in San Francisco. A major chunk of the tweets related to the announcement and participation in the event. People were inspired by the event as six women speakers (including Isis Anchalee) expressed their experiences as underrepresented minorities in the tech workforce. A number of tweets were also framed towards increasing the public awareness and highlighting the views, information and attitudes around the campaign. This aspect was expressed through the tweet terms such as: *tech, work, challenge, software,* and *share*.

Showcasing the contributions of female engineers and call for diversity in the engineering workforce also appeared as a major theme discussed and followed during the campaign. Using the hashtag #iLookLikeAnEngineer, many users (predominantly women) tweeted and posted pictures of themselves along with a description of their role in the engineering fields. These views were captured through words including *campaign, diversity, gender, challenging,* and *aims*. The hashtag originator, Isis Anchalee also received enormous support and appreciation for initiating the campaign. Women from various fields, particularly STEM encouraged other users to participate and raise public awareness by tweeting the hashtag. These sentiments were presented through terms including *female, womenintech, Twitter, great,* and *women*. Finally, the key aims of the campaign that include challenging stereotypes in engineering domain as well as erecting billboard in the Bay area also appeared as highly debated and followed topic. Tweets around this topic accompanied terms: *stereotypes, STEM, working, billboard,* and *viral*.

The hashtags analysis also reiterate a similar pattern as generated through words analysis. The campaign hashtag #Ilooklikeanenginner was obviously the leading hashtag used within the tweets.

The involvement and role of women in STEM workforce was expressed through hashtags #womenintech, #womeninstem, #women, and #Adalovelaceday. The other leading hashtags used throughout the campaign were #diversity, #race, STEM, tech, and engineering that echo the tweets calling for diversity in the STEM education and workforce.

Campaign interaction

Tweet type analysis:
A detailed analysis of the content modality of tweets indicates a very interesting pattern. Almost half of the original tweets contain text and link, followed by tweets that contain text and a photo. In total around two-third of the original tweets fall in this category, indicating the importance of photo and link in addition to tweet text. Appending link and photo point out that beside text additional information is provided that might not be possible otherwise due to text limitation. Further examination of retweets and favorites showed that tweets containing a photo stand clearly apart from other modalities.

Although only around 40% of the original tweets contain an image, majority of the highly retweeted tweets (85%) are embedded with a photo. This finding signifies the importance of embedding photos in the tweets that can have a significant impact of the retweet-ability. Likewise, all of the tweets in the most favorited list contained a photo designating a significant impact of photos on favorite-ability. This finding further endorses our inference that visual content in tweets are more appealing and perceived more positively than the other form of content or simply textual tweets by the users. Furthermore, this outcome aligns and reaffirms the power of photos in contemporary communication, especially in the context of social media where photos are used not only for social influence and disclosure but for information sharing purposes as well [18]. In the context of this campaign, an overwhelming majority of photos are of female engineers working in their workplaces or having routine daily activities e.g. spending time with family. As these pictures are closer to reality and many people can somewhat relate to them personally or someone they know. Furthermore, these picture also portray the contributions of female engineers in the workforce. Although this finding is confined to the current case, but it might be applicable to other datasets.

Network analysis:

To further understand the significant users within the community, we conducted network analysis based on the retweets. Each node represented a user (either retweeted or got retweeted) and the directed edge denoted the relationship between two nodes (users) from a user who retweets to another user. For creating the network graph and measuring various network centralities, we employed Gephi – an open source network visualization tool [19]. The #ILLAE network is composed of 24702 nodes and 27442 edges. Network visualization of the campaign is presented in Fig. 4.

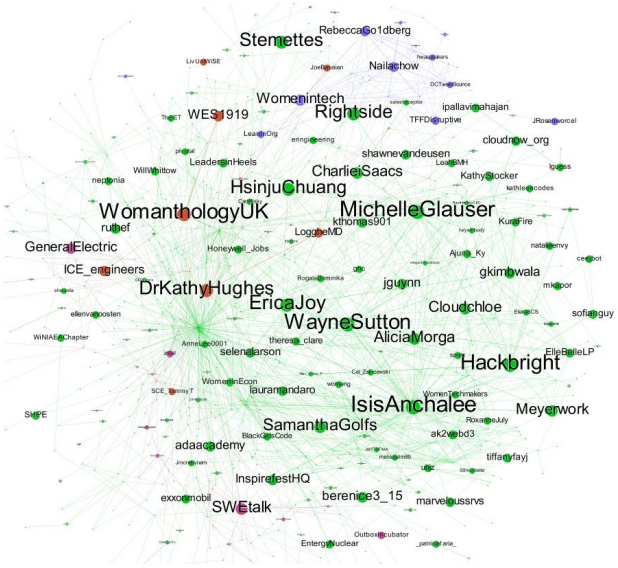


Fig. 4. Network graph of #ILLAE.

In social network analysis two key measures including degree (in-degree and out-degree) and betweenness centrality are used often to assess the key entities within the network. These metrics support in better understanding the involvement of various entities in the conversation. For instance, betweenness centrality measure helps in identifying the most influential entities within a network. The importance of a node within a network is indicated by a high betweenness centrality. On the other hand, users having a high out-degree value indicate the entities responsible for initiating most conversations (retweets in this case), meanwhile users with a high in-degree point out the participants with most of the conversations directed towards them. Node-level metrics of the campaign highlighting the top 10 entities is presented in the Table 5.

Table 5: Entities with highest out-degree, in-degree, and betweenness

Highly Gregarious		**Highly Referred**		**Highly Influential**	
Twitter handle	**Out-degree**	**Twitter handle**	**In-degree**	**Twitter handle**	**Betweeness**
MichelleGlauser	71	EngineerLook	275	IsisAnchalee	26841
IsisAnchalee	69	IsisAnchalee	110	MichelleGlauser	10917
CaterpillarInc	53	andytoronto	83	Hackbright	9869
GeneralElectric	37	liandryaaa	56	WayneSutton	8963
WayneSutton	29	SWEtalk	48	WomanthologyUK	7925
Hackbright	27	WayneSutton	42	EricaJoy	5634
TechCrunch	26	Hackbright	40	DrKathyHughes	4716
austinenergy	25	ubjz	39	Rightside	4320
hwaspeakers	23	Rosenchild	38	HsinjuChuang	4237
womeng	23	KahwaiMuriithi	30	Stemettes	4003

Analyzing the entities with high out-degree illustrates the key hubs initiating the conversations within the network. Two females including Michelle Glauser (Campaign Manager) and Isis

Anchalee (Campaign initiator) stand out as the central entities together with large engineering multinationals such as Caterpillar and General Electric. The analysis also reveal two entities including Hackbright Academy and Women in Engineering at Twitter that proactively promote the participation of women in engineering education and workforce. The in-degrees metric analysis indicates that that most of the conversational activity is directed towards the campaign organizers. The official Twitter profile of the #ILookLikeAnEngineer campaign manager and campaign initiator Isis Anchalee are most often referred in the network. Furthermore, the conversational activity was also directed towards two organizations (The Society of Women Engineers and Hackbright Academy) involved in empowerment of women in engineering and technology.

We also calculated betweenness centrality to assess the most influential actors within network. The higher the size of node the more important that actor is in connecting various communities together (see Fig. 4). Likewise, other node-level metrics, Isis Anchalee scored the highest betweenness centrality. The campaign manager Michelle Glauser, and female engineers Erica Joy, Hsin-Ju Chuang also emerged as important players in engaging people in conversation. Finally organization including Hackbright Academy, WomanthologyUK, and Stemettes challenging the stereotypes and promoting women participation in STEM were also considered as the key entities influencing the network.

Future Work

In our future work we plan to work with larger datasets and apply techniques such as supervised classification [20], [21] to help us categorize a large set of users automatically. For analyzing interaction patterns of the categorized users, we will create interaction networks [22] by considering nodes as users and edges as the interaction strength—who spreads whose messages, who acknowledges whom, etc. We will use different types of centrality measures and network visualization algorithms, such as Force-Directed layout those available in popular open source tools like Gephi, to identify patterns of clusters, and influential users. This technique will be useful in identifying and understanding interactions of the actors – individuals and organizations – within our dataset that participate in STEM related issues. Topic modeling is a content analysis method to identify hidden topics in the content of discourse and a topic model provides topics as word clusters where the clustered words represent semantically related meaning. For instance, a topic of computer engineering in the discourse might be expressed by use of several engineering related words in different tweet messages—computer, software architecture, scalable computing, cloud backend, system design, development process, etc. We will ultimately use the topical distributions of such word clusters over messages to automatically partition dataset by different topics, and it will help us understand dynamics of participation in the discourse related to diverse STEM issues. Sentiment-emotion mining is a text analysis method to understand subjective behavior of the text author. Among the different variants of the method, a relevant technique to our research is target-specific sentiment estimation, which measures sentiment for a given target in the text. For example, a target can be specific STEM issue of ‘women’, and therefore, whenever a Twitter message contains words related to the target ‘women’, this technique provides the valence measure of the message author towards women as positive, negative or neutral with probability. This group of techniques will help us gauge the attitude of participating actors in STEM related discourse towards specific issues. Through this we will be able to develop a deeper picture of what participants feel about STEM.

Conclusion

We analyzed data from a hashtag activism campaign, #ILookLikeAnEngineer, to identify how the affordances of social media, in our case Twitter, support social movements for STEM diversity. The ability of a diverse range of people to participate in a campaign and to voice their opinions is critical for a broad support to an issue of social good. We find that in the case of this campaign Twitter allowed users to express themselves in different ways but implicitly related, and participate regardless of whether they were an individual or a corporation lending the campaign a strong support and momentum. Our empirical analysis also sheds lights on issues that people consider to be important in order to advance gender diversity within the engineering/technology workplace. Unlike basic observations of the activity of user engagement using counts, a key outcome of such an analysis will be better understanding of patterns of how campaigns and specific issue-based hashtags generate varied types of engagement from the public and where.Overall, our preliminary work provides enough evidence that looking at social media data, Twitter data in our case, provides useful information and when aggregated and thoroughly parsed and analyzed, this information has the potential to provide novel insights related to STEM diversity in education and workforce. We developed the analytical foci for the proposed research after the preliminary work. The preliminary research questions were small in scope but they helped us sharpen our analytical foci to provide a more nuance and useful look at the data.

References

[1] M. Sanders, "STEM, STEM Education, STEMmania.," *Technology Teacher*, vol. 68, no. 4, pp. 20–26, 2009.

[2] National Research Council, "Rising above the gathering storm, revisited: Rapidly approaching category 5," The National Academies Press, Washington, DC, 2010.

[3] H. B. Gonzalez and J. J. Kuenzi, "Science, technology, engineering, and mathematics (STEM) education: A primer," 2012.

[4] M. Feuer, "STEM education: Progress and prospects," *The Bridge*, vol. 43, no. 1, pp. 3–6, 2013.

[5] L. E. Malcom-Piqueux and S. M. Malcom, "Engineering diversity: Fixing the educational system to promote equity," *Bridge*, vol. 43, no. 1, pp. 24–34, 2013.

[6] A. L. Kotsenas, M. Arce, L. Aase, F. K. Timimi, C. Young, and J. T. Wald, "The strategic imperative for the use of social media in health care," *Journal of the American College of Radiology*, vol. 15, no. 1, pp. 155–161, 2018.

[7] T. Highfield, *Social media and everyday politics*. John Wiley & Sons, 2017.

[8] C. Reuter and M.-A. Kaufhold, "Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics," *Journal of Contingencies and Crisis Management*, 2017.

[9] R. Rickford, "Black lives matter: Toward a modern practice of mass struggle," in *New Labor Forum*, 2016, vol. 25, pp. 34–42.

[10] M. Hänska Ahy, "Networked communication and the Arab Spring: Linking broadcast and social media," *New Media & Society*, vol. 18, no. 1, pp. 99–116, 2016.

[11] A. Johri, H. Karbasian, A. Malik, R. Handa, and H. Purohit, "How Diverse Users and Activities Trigger Connective Action via Social Media: Lessons from the Twitter Hashtag Campaign# ILookLikeAnEngineer," in *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii, USA, 2018.

[12] H. Achrekar, A. Gandhe, R. Lazarus, S.-H. Yu, and B. Liu, “Predicting flu trends using twitter data,” in *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, 2011, pp. 702–707.
[13] A. Jolly, “Finding Real-World Problems for STEM Lessons.” [Online]. Available: https://www.middleweb.com/5003/real-world-stem-problems/. [Accessed: 09-Mar-2018].
[14] C. C. Seepersad, “How to involve more women and girls in engineering,” *The Conversation*. [Online]. Available: http://theconversation.com/how-to-involve-more-women-and-girls-in-engineering-55794. [Accessed: 09-Mar-2018].
[15] D. Miller, “Men and women biased about studies of STEM gender bias – in opposite directions,” *The Conversation*. [Online]. Available: http://theconversation.com/men-and-women-biased-about-studies-of-stem-gender-bias-in-opposite-directions-48924. [Accessed: 09-Mar-2018].
[16] D. Lazer *et al.*, “Life in the network: the coming age of computational social science,” *Science (New York, NY)*, vol. 323, no. 5915, p. 721, 2009.
[17] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.
[18] A. Malik, A. Dhir, and M. Nieminen, “Uses and gratifications of digital photo sharing on Facebook,” *Telematics and Informatics*, vol. 33, no. 1, pp. 129–138, 2016.
[19] M. Bastian, S. Heymann, and M. Jacomy, “Gephi: an open source software for exploring and manipulating networks,” *ICWSM*, vol. 8, pp. 361–362, 2009.
[20] H. Karbasian, H. Purohit, R. Handa, A. Malik, and A. Johri, “Real-Time Inference of User Types to Assist with more Inclusive and Diverse Social Media Activism Campaigns,” in *Proceedings of the 1st AAAI/ACM Conference on AI, Ethics, and Society*, New Orleans, USA, 2018.
[21] M. Pennacchiotti and A.-M. Popescu, “A Machine Learning Approach to Twitter User Classification,” *ICWSM*, vol. 11, no. 1, pp. 281–288, 2011.
[22] H. Purohit, J. Ajmera, S. Joshi, A. Verma, and A. P. Sheth, “Finding Influential Authors in Brand-Page Communities,” in *ICWSM*, 2012.

Computers and Electrical Engineering 68 (2018) 412–424

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

Improved diagnosis of Parkinson's disease using optimized crow search algorithm☆

Deepak Gupta[a], Shirsh Sundaram[a], Ashish Khanna[a], Aboul Ella Hassanien[b], Victor Hugo C. de Albuquerque[c,*]

[a] Maharaja Agrasen Institute of Technology, Delhi, India
[b] Faculty of Computers and Information, Cairo University, Cairo, Egypt
[c] Graduate Program in Applied Informatics, University of Fortaleza, Fortaleza, CE, Brazil

ARTICLE INFO

Keywords:
Parkinson's disease
Crow search algorithm
Optimized crow search algorithm
Feature selection
Chaotic crow search algorithm
Decision Tree
Random Forest
k-Nearest Neighbor

ABSTRACT

Diagnosis of Parkinson's disease at its early stage is important in proper treatment of the patients so they can lead productive lives for as long as possible. Although many techniques have been proposed to diagnose the Parkinson's disease at an early stage but none of them are efficient. In this work, to improve the diagnosis of Parkinson's disease, we have introduced a novel improved and optimized version of crow search algorithm(OCSA). The proposed OCSA can be used in predicting the Parkinson's disease with an accuracy of 100% and help individual to have proper treatment at early stage. The performance of OCSA has been measured for 20 benchmark datasets and the results have been compared with the original chaotic crow search algorithm(CCSA). The experimental result reveals that the proposed nature-inspired algorithm finds an optimal subset of features, maximizing the accuracy and minimizing a number of features selected and is more stable.

1. Introduction

Parkinson's disease (PD), first described by James Parkinson, affects one's movement which are changes in writing and speech skills, tremors, stiffness in muscles. This happens when the nerve cells which produce an important chemical called dopamine breaks down making nerve cells unable to send messages properly. This may happen due to hereditary reasons. This causes memory impairment, nervous system disorders, depressions. There's been a lot of research done for diagnosing the disease at its early stage, though not with reliable success. It is important to diagnose the disease at its early stage so the patients can live quality lives for as long as possible. The disease at its later stages affects the daily tasks and the patient may need assistance from others. The advanced stages of Parkinson's disease are more dangerous as the patients get advanced stiffness in the legs that can also cause freezing upon standing, making it impossible to stand or walk. In the recent times, many approaches have been used to identify the disease correctly like voice and speech exams, writing exams. The handwriting exam is used widely for identification of Parkinson's disease since it is cheaper and easier to get data. The work of Rosenblum [1] proposes that writing exams can be used to differentiate between a PD patient and healthy individuals. Recently, Pereira [2] conducted work to extract features from writing exams using image processing techniques and prepared a dataset called "HandPD" with all images and features extracted. Identification of Parkinson's disease using

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. G. R. Gonzalez.

* Corresponding author.
E-mail addresses: deepakgupta@mait.ac.in (D. Gupta), ashishkhanna@mait.ac.in (A. Khanna), aboitcairo@fci-cu.edu.eg (A. Ella Hassanien), victor.albuquerque@unifor.br (V.H.C. de Albuquerque).

https://doi.org/10.1016/j.compeleceng.2018.04.014
Received 8 March 2018; Received in revised form 20 April 2018; Accepted 23 April 2018
0045-7906/ © 2018 Elsevier Ltd. All rights reserved.

deep optimum path forest [3]. Using the Restricted Boltzmann machine for diagnosis of Parkinson's disease from handwritten dynamics and firefly algorithm is used for pattern recognition. In this paper, we have used optimized crow search algorithm for identification of Parkinson's disease in patients so they can receive proper treatment and live quality lives. The proposed algorithm OCSA is based on the evolutionary algorithm CSA [4] which uses intelligent behaviors of crows to find the best solutions to problems.

An evolutionary computation has attracted great attention for solving optimization problems in the recent years. Evolutionary algorithms are the generic population based bio-inspired optimization algorithm which mimics biological evolution such as mutation, reproduction, selection to solve optimization problems. Many algorithms have been proposed by taking inspirations from nature in recent studies, such as genetic algorithms (GA) are based on natural selection [5], bat algorithm has been inspired by echolocation behaviors of microbats [6], particle swarm optimization uses intelligent collective behavior of birds flocks and fish schools [7], crow search algorithm (CSA) is another metaheuristic approach inspired by the intelligent behaviors of crows while hiding and stealing foods, it is believed that crows are among the world's most intelligent animals [8,9]. Recent application of evolutionary algorithms are: Ant Colony Optimization based clustering for analysis of Ego Network [10] which proposes a model to find and analyze the circles in the social network using ant colony optimization, Mission planning problems can be solved using a knee point based MOGA(multi-objective genetic algorithm) [11], Complex multi-UAV mission planning problems can be solved using (MOGA) [12], Boosting process, guided by Genetic Algorithm (GA), is used for malware detection [13], Evolving deep neural networks architectures for Android malware classification [14], Evolutionary multi-objective classifier is created for malware detection in Android [15], A study is done on combined intelligence algorithms for solving problems of video games [16], Ants colonies are used differently for simulations of micro and macro lemmings [17], Multiple classifiers in Biometrics: Fundamentals and review, trends and challenges are discussed in the paper referenced in [18,19]. Though EAs have different structures they are similar to most of EAs work on random distribution in the search space. EAs need to provide a proper balance between exploration and exploitation [20], otherwise, algorithms usually get stuck in the local minima.

In the recent years, data have been increased in the form of numbers of features and number of instances, which makes data noisier. The noisier datasets can cause the models to increase computational costs, increase complexity, training the data slower, decrease the accuracy predicted by models. Hence feature selection has become an important task for machine learning before training the model. Feature selection, aka attribute selection, is a technique that aims to find a subset from the given complete set of feature and fewer downgrades of the performances of the system; so that the subset of features predicts the target with accuracy comparable to the performance of the original set of features, and with the lessening of computational cost. Feature selection can be mainly categorized as filter-based and wrapper-based algorithm. The wrapper-based algorithm, computational costlier than the filter method, uses machine learning algorithm for evaluation, and filter-based algorithm uses a statistical approach for scoring each feature. Evolutionary algorithms can be modified; such that they can be used for feature selection, such as Binary Bat Algorithm [21], Modified Binary Bat Algorithm [22], and Binary chaotic crow search algorithm [24].

The main contributions of this work are:

1. A novel natural inspired optimization algorithm named optimized crow search algorithm (OCSA) for efficient feature selection has been introduced.
2. OCSA is evaluated on HandPD and NewHandPD datasets for the identification of Parkinson's disease with the accuracy of 100%.
3. The proposed algorithm OCSA is validated on 20 benchmark datasets to select the best subset of features from an original set of feature of datasets without downgrading the performance and the algorithm is compared with chaotic crow search algorithm (CCSA). The average accuracy (88.2%) and computational time (0.20 s) of OCSA are enhancing the performance of CCSA with an accuracy of 84.2% and computational time of 0.40 s.
4. We would also like to highlight that the main goal of proposing and implementing the optimized crow search algorithm is to keep it as easy, simple and understandable as possible.
5. For classification purposes, we used three different algorithms: (i) k-nearest neighbors (k-NN), (ii) Random Forest and (iii) Decision Tree classifiers. The results show that the Random Forest algorithm used in OCSA outperforms other machine learning models with a significant boost to the performance of CCSA.

In the following Section 2, the preliminary algorithms and techniques have been presented; Section 3 introduces OCSA for feature selection followed by the implementation of proposed OCSA for feature selection in Section 4. Results are discussed in Section 5; followed, by the conclusion and future perspectives.

2. Background

2.1. Crow search algorithm

In 2016, Askarzadeh proposed a metaheuristic algorithm called crow search algorithm [4]. It is inspired by the intelligent behaviors of crows while hiding and stealing foods. It is now believed that crows are among the world's most intelligent animals [8,9]. There are many pieces of evidence showing crows' cleverness. Crows are self-aware as shown in the mirror test. Crows can remember faces, exchange information with each other, steal the food and keep it away from others by hiding the collection of foods [8,24]. They are intelligent thieves as they take extra care such as changing hiding places time to time so that they can avoid being victims in the future. Crows use a knowledgeable approach to hiding their food and stealing other's food.

Using above clever behaviors of the crows, Askarzadeh proposed crow search algorithm. The concept behind crow search

Algorithm 1
Crow search algorithm.

Input: Constrained Engineering design problems having objective functions, decision variables and constraints.

Output: Optimized solution for the Objective functions.

Auxiliary:

```
Set the values of M, Ap, fl, tmax, D
Initialize the crow position x and the memory of crows with random values
for t in range ( tmax ):
    for i in range ( N ):
        Select a crow randomly to follow
        Set the value of r_i randomly
        if r_i >= Ap :
            x_{i,t+1} ← x_{i,t} + r_i * fl * ( mem_{i,t} - x_{i,t} )
        else :
            x_{i,t+1} ← a random position.
        Check how feasible value of x is
    Calculate the new position x_{i,t+1} of each crow
    Update the memory of each crow
```

algorithm are: crows stay in the flock, crows keep hiding places of food in memory, crows go after each other to steal food, and to prevent stealing crows protect their collections by using probability.

2.1.1. CSA implementation of optimization problems

Let *M* be the number of crows, *Ap* is awareness probability, *D* represents the total number of dimensions, *fl* is the flight length, the variable *tmax* defines a number of iterations. The memory of crows are represented as, *mem*, to store the hiding places of the food. The variable *Ap* makes sure the balance is maintained between the exploring and exploiting; D is same as the number of features.

Now, each crow goes after a random crow from the flock and then two cases can arise. The first case arises when crow does not know that another crow is following it. The Second case arises when the crow comes to know that another crow is going after it.

Case 1: The crow is not aware of being followed,

$$x_{i,t+1} = x_{i,t} + r_i * fl * (mem_{i,t} - x_{i,t}), \quad if\ r_i >= Ap, \tag{1}$$

where, x is the updated position of the crows, r is a randomly selected number taking values between [0,1].

Case 2: The crow is aware of being followed, in which $x_{i,t+1} = a\ random\ position.$

The crow search algorithm is defined as (Algorithm 1):

2.2. Chaotic crow search algorithm

Sayed et al. proposed a novel bio-inspired optimizer, called by chaotic crow search algorithm [23], which is based on CSA to optimize the features selection problems. The algorithm is applied to 20 different datasets. CCSA overcomes the problems posed by the CSA and many others metaheuristic algorithms. They have proposed a hybrid approach by using the chaotic search method for selecting the optimal feature subset. CCSA have used a wrapper based approach for feature selection. A fitness function is introduced which uses the wrapper based approach as for selecting the optimal features. The fitness function is given by $ft = acc + wf * (1 - lf / lt)$, where acc is accuracy predicted by the model k-Nearest Neighbor when applied on the feature subset of length *lf*. *lt* is a number of features in the datasets.

The basic steps involved in this algorithm are as given below (Algorithm 2):

2.3. Features selection

Features selection is done to decrease the number of features from the dataset in such a way that the performance of the model remains same or increases in the selected features. The target of the algorithms is to find a subset from the complete set of features so that the subset of features predict the target with accuracy comparable to the performance of the original set of features, and with the lessening of computational cost. Feature selection is used due to the following reasons (1) it reduces the complexity of a model, (2) it enables the algorithm to train faster, (3) it reduces overfitting and (4) it improves the accuracy of a model if the right subset is chosen.

Algorithm 2
Chaotic crow search algorithm.

Input: Different datasets with a different set of features associated with each crow replacing basic randomness with chaos.

Output: Subset of the feature, and accuracy using the chaotic binary version of crow search algorithm.

1. *Set the values of M, Ap, fl, tmax, D*
2. *Initialize the crow position x randomly*
3. *Calculate the fitness function, ft, of each crow*
4. *Initialize the memory of crows mem*
5. **for** *t* **in range (** *tmax* **) :**
6. **for** *i* **in range (** *N* **) :**
7. *Select a crow randomly to follow*
8. *Set the value of* r_i *randomly*
9. **if** $r_i >= Ap$ **then**
10. $x_{i,t+1} \leftarrow x_{i,t} + r_i * fl * (mem_{i,t} - x_{i,t})$
11. **else**
12. $x_{i,t+1} \leftarrow$ *a random position.*
13. *Convert the position into binary using a sigmoid function*
14. *Check how feasible value of x is*
15. *Calculate the new position* $x_{i,t+1}$ *of each crow*
16. *Update the memory of each crow*
17. *Output the best solution*

3. Proposed optimized crow search algorithm (OCSA)

The proposed *OCSA* algorithm is implemented, using Python language program, as a feature selection algorithm-based filter method, selecting the most optimal subset of features from the dataset *OCSA*, following functions:

Init function: This function initializes the position of the crows in the space. Every position of each crow is initialized with the value 1.

Fitness function: fitness function is used for updating the positions of the crow in the search dimension. The fitness is calculated for each crow by the following formula given in Eq. (2).

$$fnt = imp + wf \times \left(1 - \frac{lf}{lt}\right) \tag{2}$$

where *fnt* is the fitness of each crow, *imp* is the importance of features calculated using *feature_importances_,* an attribute of the Random Forest classifier [3], *wf* is the weight factor used in the fitness function to control the balance between features importance and number of features, *lf* is the number of attributes selected, and *lt* is total number of attributes. A Random Forest is an ensemble learning method that develops large numbers of random Decision Trees analyzing sets of variables and it uses averaging to improve predictive accuracy and control over-fitting. feature_importances_ returns the importance of each feature. The higher the value, the more importance the feature.

The proposed algorithm of *OCSA* is described below:

Following changes has been proposed in the optimized crow search algorithm as compared to original chaotic crow search algorithm:

- In lines 1–4, the list "imp" returns the feature importance of each feature. Higher the value, higher the importance.
- In line 5, the positions of crow are initialized with the value of 1, so each feature is available for selection at the initial stage.
- The fitness function generates the fitness of each crow. The fitness is calculating the fitness using the feature importance of every feature and the length of feature subset using the position xn of each crow. wf is the weight factor in fitness function is changed such that it controls the balance between features importance and number of features.
- In line 9, the random function returns a random value and store it in r.
- In lines 22–28, the feasibility of each crow is tested. If the crow position has at least one position value equal to 1 and new fitness generated is greater than the old one the positions of the crow will be updated. Else all the position of the crow is made equal to 1.
- In lines 29–32, after getting the best feature subset, it is trained by the techniques: k-NN, Random Forest, Decision Tree and the accuracy is predicted. The feature subset and the accuracy are given as output.

Table 1
Input parameters.

Parameters	Values	Description
m	30	Number of crows
Ap	0.1	Awareness probability
fl	0.2	Flight length
tmax	50	Number of iterations
d	Same as number of features	Number of features
Wf[a]	1.5	Weight factor

[a] Weight factor: It is a control parameter used in the fitness function to control the balance between features importance and number of features selected.

4. Implementation of the proposed algorithm

In this section, the proposed algorithm has been implemented. The experimental setup, initial input parameters, datasets used and the implementation have been discussed.

4.1. Experimental setup

To test the proposed algorithm, a desktop personal computer having 1.80 GHz Core 2 CPU and minimum 2GB of memory under Ubuntu version 14.04 is used. We implemented the algorithms using Python version 3.6.3::Anaconda, Inc.

4.2. Input parameters

The parameters are initialized at the beginning of the algorithm as given in Table 1.

4.3. Datasets

OCSA is applied to Parkinson's disease datasets and then to the 20 benchmark datasets. We have discussed the datasets in details.

4.3.1. Parkinson's disease

The algorithm is applied to the HandPD dataset of Parkinson's disease as described in the paper of Pereira. PD affects the ability of the patients to move and the disease becomes more dangerous as it advances to its later stages. So it is important to diagnose the disease at early stages . Pereira [2] prepared datasets by extracting features from handwritten exams taken by the patient group and control group. Though there are many different approaches to conduct examination like voice and speech exams, writing exams but handwritten exams are selected since it is cheap and easier to get. The dataset is prepared by taking handwritten exams from the healthy group and the patients group who have Parkinson's disease (PD). The handwritten exams were prepared at Botucatu Medical School, São Paulo State University - Brazil. The exam consists of a form which has four spirals and four meanders as shown in the below Figs. 1 and 2 [2]. The following images show the examples of the HandPD dataset.

The exam is taken by 92 individuals of which 18 are healthy individuals and 74 are patients. Healthy group comprised of 6 male and 12 female individuals with average age of 44.22 ± 16.53 years. Among those individuals, 2 are left-handed and 16 are right-handed. Patient groups comprised of 59 male and 15 female individuals with average age of 58.75 ± 7.51 years. Among those individuals, 5 are left-handed and 69 are right-handed.

A NewHandPD dataset is also made available. The exam is taken by 66 individuals of which 35 individuals are healthy and 31 are

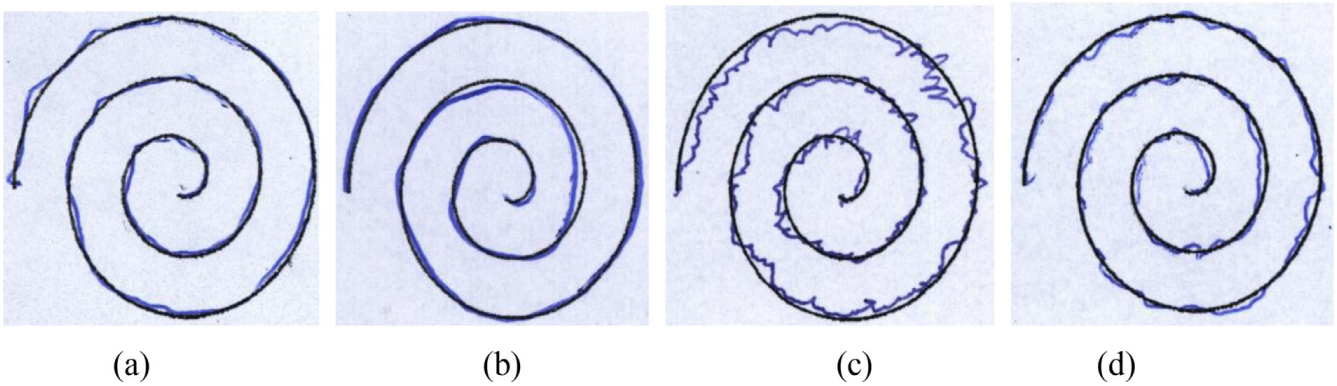


Fig. 1. Examples of spirals from HandPD dataset: (a) 58-years old male and (b) 28-years old female individuals of control group, and (c) 56-years old male and (d) 65-years old female individuals of patient group.

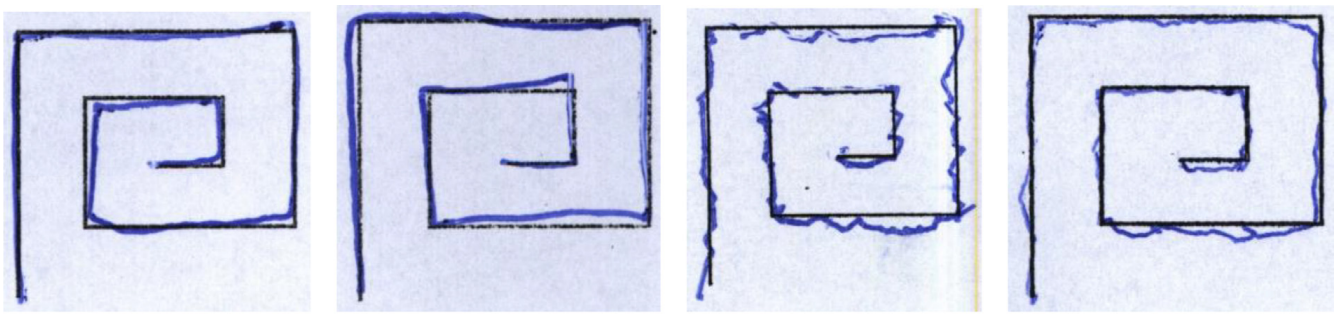


Fig. 2. Examples of meanders from HandPD dataset: (a) 58-years old male and (b) 28-years old female individuals of control group, and (c) 56-years old male and (d) 65-years old female individuals of patient group.

Table 2
Features descriptions of the HandPD and NewHandPD datasets.

Name of the feature	Descriptions
CLASS_TYPE	1 = control group and 2 = patient group
ID_EXAM	Identifier of the exam
IMAGE_NAME	Name of the image
ID_PATIENT	Patient id
GENDER	M = male and F = female
RIGHT/LEFT-HANDED	R = right handed and L = left handed
AGE	Age (years) of the patient
RMS	Root mean square of the difference between HT(handwritten trace) and ET (exam template) radius
MAX_BETWEEN_ET_HT	The maximum difference between ET (exam template) and HT (handwritten trace) radius
MIN_BETWEEN_ET_HT	The maximum difference between ET and HT radius
STD_DEVIATION_ET_HT	Standard deviation of the difference between ET and HT radius
MRT	Mean relative tremor: the mean difference between the radius of a given sample and its d left-nearest neighbors
MAX_HT	Maximum HT radius
MIN_HT	Minimum HT radius
STD_HT	Standard deviation of HT radius
CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT	The number of times the difference between ET and HT radius changes from negative to positive, or vice-versa

patients. Each patient was asked to draw 4 spirals 4 meanders 2 circles (one in the air and one in paper), and left and right-handed diadochokinesis. The feature description of HandPD and NewHandPD datasets have been given in Table 2.

Here ET stands for Exam Template that is the actual figure (spiral or meander) in the form and HT stands for Handwritten trace (spiral or meander) drawn by the individual as shown in Figs. 1 and 2 respectively.

4.3.2. 20 Benchmark datasets

For benchmark testing, we have used 20 datasets of varying types including medical and business data sets. The data sets are collected from the UC Irvine machine learning repository [25]. The datasets have been described in Table 3. The reason behind selecting these 20 datasets is testing whether the proposed algorithm can be evaluated on different types of datasets or not. Different types of datasets are Game, Business, Physical, Life, Computer, Clinical. Here some of the datasets contain missing values. We have used the most occurring value in the feature to replace the missing values that are the mode of each feature is calculated. Then mode is used for replacing all the missing values in the respective features.

4.4. Implementation

The binary crow search algorithm is modified to improve the accuracy with reduced features and reduced computational costs. The fitness function is changed as it is calculating feature importance instead of predicting the accuracy using some model which increases the computational cost. Three different machine learning techniques viz. k-Nearest Neighbors, Random Forest, Decision Trees are used for predicting the accuracy. The results obtained are compared with the results from *CCSA*. Holdout method is considered with a split of 75% of the dataset for training model and 25% of the dataset for testing.

Table 3
20 datasets from UCI machine learning repository.

Id	Datasets	No of classes	Missing values	No. of features	No of instances	Type
1	Chess	2	No	36	3196	Game
2	Poker hand	10	No	10	25,010	Game
3	German credit	2	No	24	1000	Business
4	Credit approval	2	Yes	15	690	Business
5	Cylinder bands	2	Yes	40	512	Physical
6	Abalone	29	No	8	4177	Life
7	Glass identification	6	No	10	214	Physical
8	Letter recognition	26	No	17	20,000	Computer
9	Waveform	3	No	21	5000	Physical
10	Zoo	2	No	18	101	Life
11	Wisconsin diagnosis breast cancer	2	No	325	596	Clinical
12	Mice Protein Expression (MPED)	8	Yes	82	1080	Clinical
13	Detection of Parkinson's disease	2	No	23	197	Clinical
14	Cardiotocography	3	No	23	2126	Clinical
15	Hepatitis	2	Yes	19	155	Clinical
16	Lung cancer	3	Yes	56	32	Clinical
17	Single Proton Emission Computed Tomography (SPECT)	2	No	44	267	Clinical
18	Thoracic surgery dataset	2	No	17	470	Clinical
19	Statlog (heart)	2	No	13	270	Clinical
20	Indian liver patient dataset	2	No	10	583	Clinical

Table 4
Configuration of the algorithms adopted.

Model	Tuning parameters
Decision Trees	max_depth = 30 min_sample_split = 20
Random Forest	n_estimators = 500
k-NN	n_neighbors = 13

Table 5
Features selected by OCSA.

S. No.	Name of the feature selected
1	GENDER
2	AGE
3	RMS
4	MAX_BETWEEN_ET_HT
5	MIN_BETWEEN_ET_HT
6	MAX_HT
7	CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT

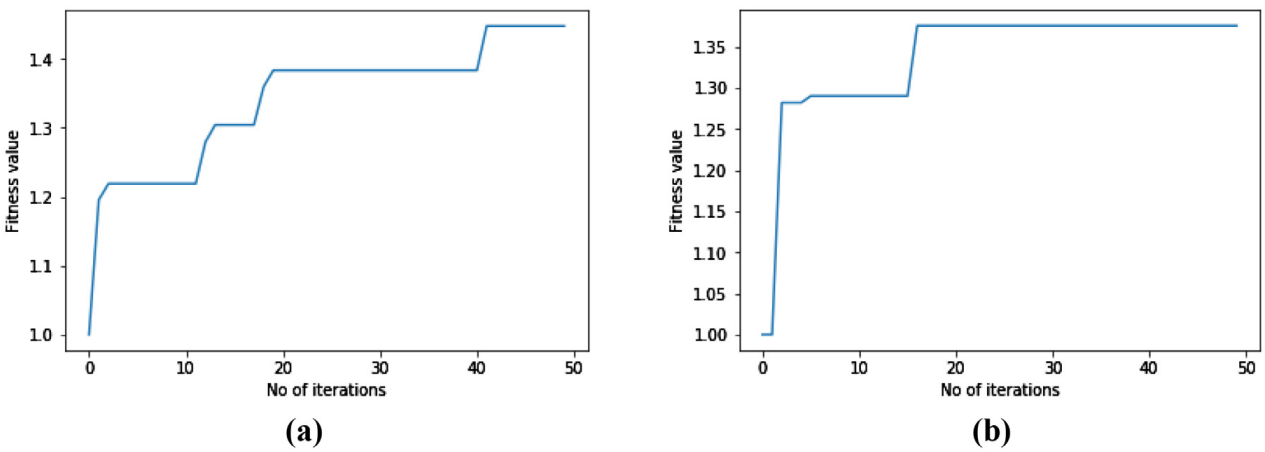


Fig. 3. Plots of convergence rate calculated on (a) spiral and (b) meander datasets respectively.

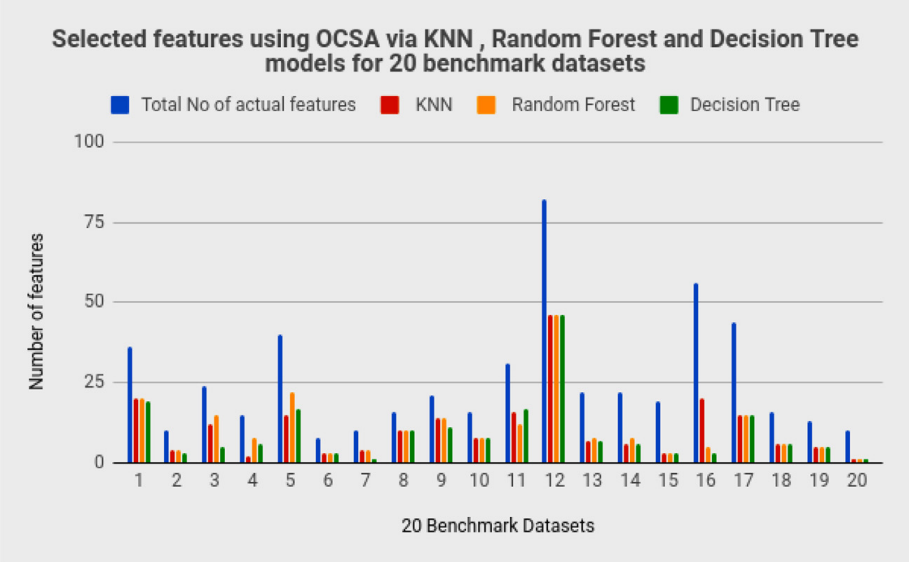


Fig. 4. Plot of selected features for 20 benchmark datasets using OCSA for different models.

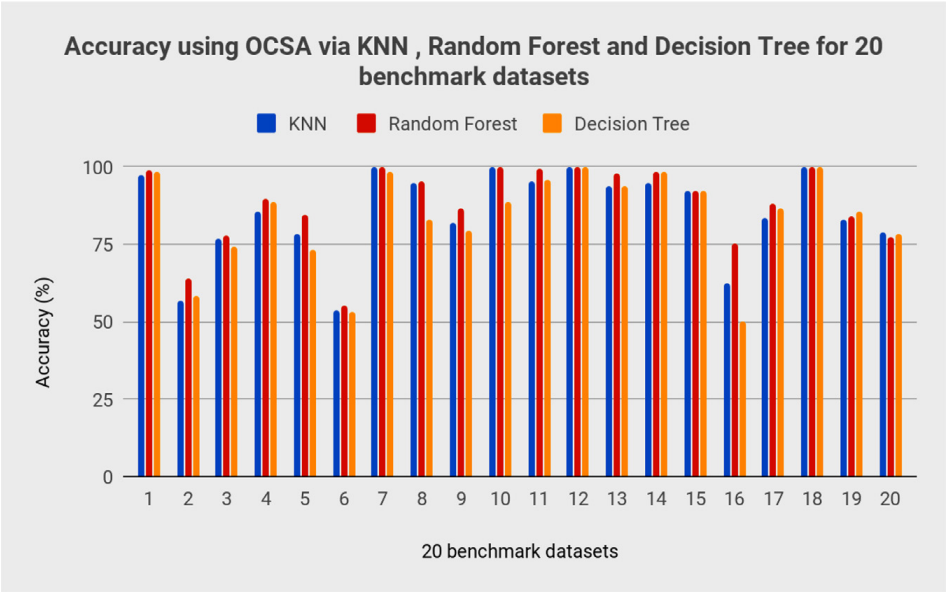


Fig. 5. Plot of accuracy for 20 benchmark datasets using OCSA for different models.

4.4.1. Machine learning models

We have used three well-known machine learning techniques for evaluating the efficient of proposed feature selection method, such as k-Nearest Neighbors (k-NN), Random Forest and Decision Trees. k-NN is the supervised learning algorithm used for classification and regression problems. Here k is a positive, typically small integer; kNN categorizes the object to the class which is most common among the nearest k neighbors. The value of k is set to be $k = 3$. Random Forest uses ensemble learning methods for solving classification and regression problems. Random Forest creates a forest with some trees, higher the number of trees higher the accuracy. We have set a number of trees equal to 500 as you can see in the table below. Decision Trees are the supervised based learning algorithms, which can be mostly used in the classification problem. It split the total population into two or more sub-population based on the input variable's most significant splitter. Table 4 shows the basic tuning parameters of the machine learning models.

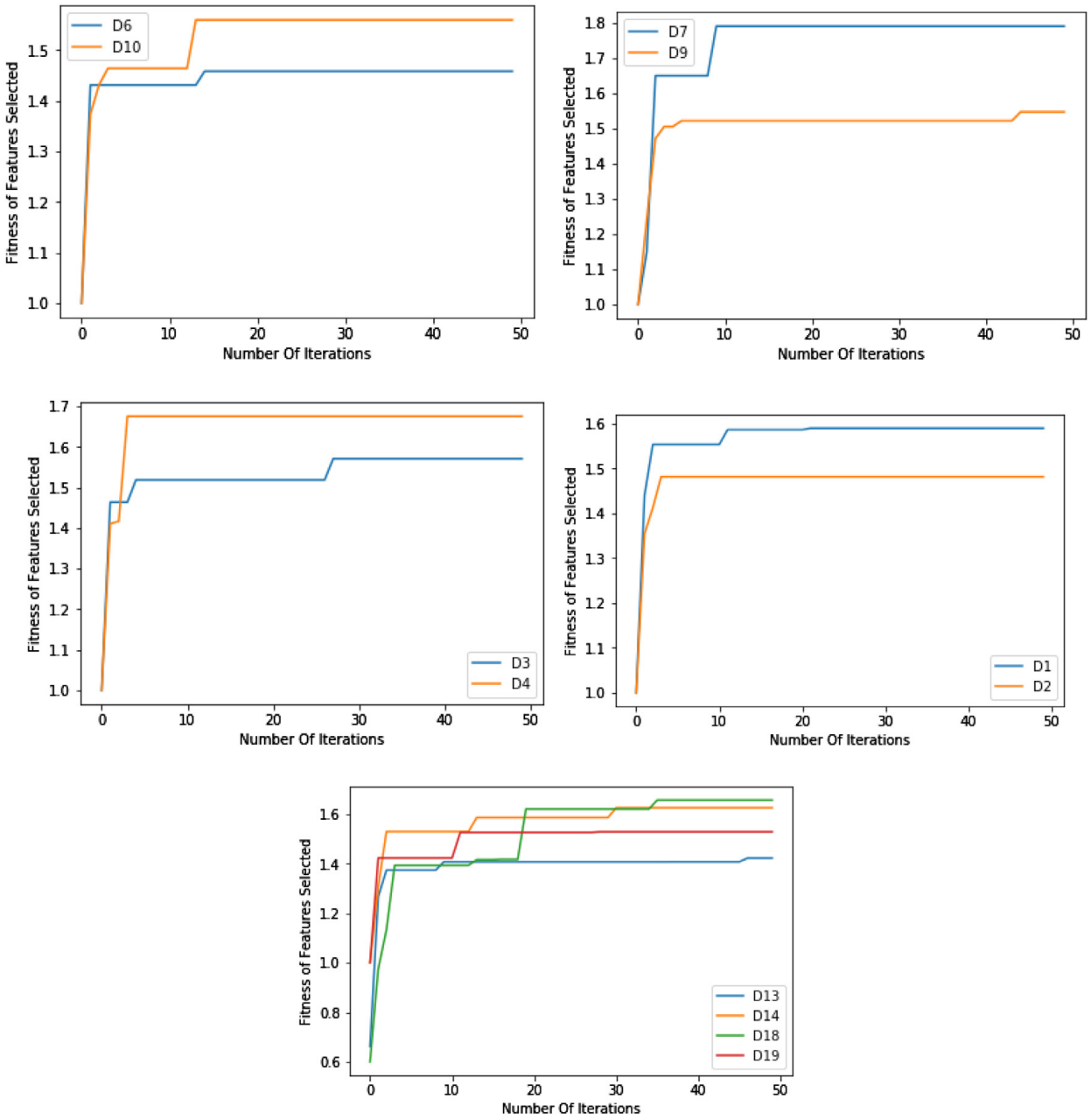


Fig. 6. Plots of Convergence rate calculated on the dataset by grouping them into five groups.

5. Results and discussions

This section discusses the results generated when the proposed optimized crow search algorithm have been applied to Parkinson's disease dataset and 20 benchmark datasets.

5.1. Dataset of Parkinson's disease

The HandPD dataset and NewHandPD dataset are applied to the OCSA algorithm. The algorithm selects only the 7 features out of the original set of features of the datasets as seen in Table 5. The accuracy over 7 selected features by the OCSA algorithm is 100% for the datasets HandPD and NewHandPD. All the three machine learning classifiers (Random Forest, Decision Tree, k-Nearest Neighbor) used by OCSA algorithm predicts 100% accuracy. The convergence rate is defined as how fast the algorithm converges towards the

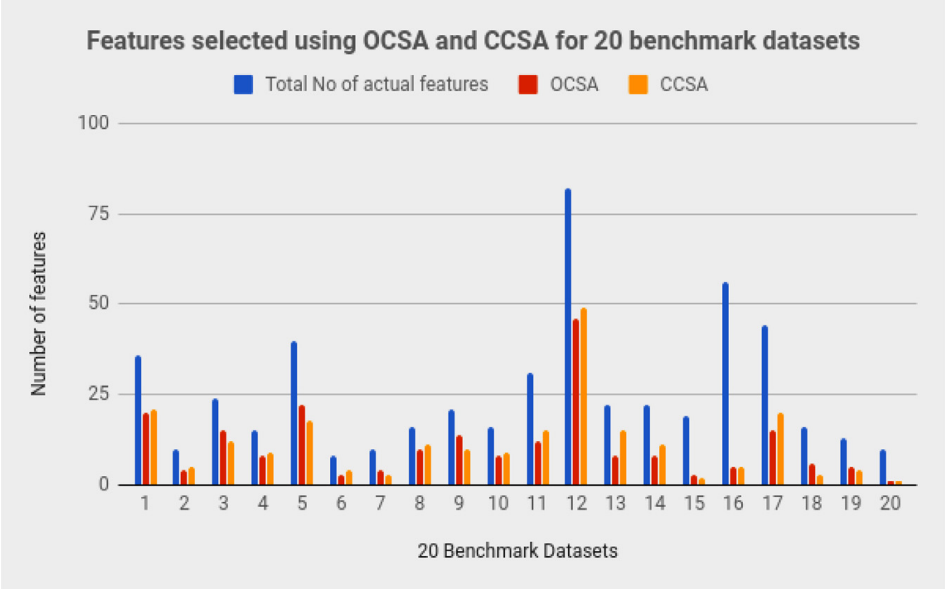


Fig. 7. Plot of selected features of 20 benchmark datasets for OCSA and CCSA.

global optima. The convergence rate for Spiral and Meander dataset have been presented in Fig. 3(a) and (b) respectively.

As seen from the results, the proposed algorithm has successfully predicting whether the patients are suffering from Parkinson's disease or not; thus this model can be used for predicting the disease and help them to have proper treatment at an early stage.

5.2. 20 benchmark datasets

The benchmark datasets given in Table 3 have been applied to the *OCSA* algorithm via K-Nearest Neighbor, Random Forest, and Decision Tree classifiers to obtain a set of features and to compute mean accuracy we conducted a cross-validation over ten runs. The plot of selected attributes using the proposed algorithm is presented in Fig. 4. The accuracy over selected attributes using different classifiers is presented in Fig. 5.

The result shows that the proposed *OCSA* algorithm selects fewer numbers of features, generating higher average accuracy values of 88.2% and computational cost of 0.20 s and fewer downgrades the performances of the system. The novel optimized crow search optimization algorithm has been generating efficient and better results if it is used with Random Forest model.

Convergence rate for the 20 benchmark datasets has been found as given in Fig. 6. Convergence rate has been calculated by grouping the datasets into five groups viz Life (D6 and D10), Physical (D7 and D9), Business (D3 and D4), Game (D1 and D2), and Clinical (D13, D14, D18, D19). The convergence curve shows that the proposed algorithm is stable.

The fitness value reaches its optimum value at the 15th and 14th iteration for dataset D6 and D10 respectively, 10th and 43rd iteration for dataset D7 and D9 respectively, 28th and 5th iteration for dataset D3 and D4 respectively, 22nd and 5th iteration for dataset D6 and D10 respectively, 46th, 31st, 36th and 29th iteration for dataset D13, D14, D18 and D19 respectively.

Thus, in this section, we have optimized the crow search algorithm, which accepts the dataset cum training set as an input and generates the optimal reduced set of features out of all the features and comparative accuracy without degrading the performance of the system as an output. We have also calculated the convergence rates of the datasets for the stability of fitness function of the proposed algorithm.

we have compared the results of proposed *OCSA* algorithm with the original *CCSA* for 20 benchmark datasets selected from UCI repository. The plots for the reduced set of features, accuracy and computational time using proposed OCSA and original CCSA have been shown in Figs. 7–9 respectively.

The proposed *OCSA* optimization algorithm outperforms the original *CCSA* for most of the datasets, as it selects fewer numbers of features, generating higher accuracy values of 88.2%, computational cost of 0.20 s and fewer downgrades the performances of the system as compared to average accuracy of 84.2% and computational cost of 0.40 s of the original *CCSA*.

6. conclusions and future works

A novel modified version of the naturally inspired optimization algorithm for feature selection named *optimized crow search algorithm* has been proposed in this work to get reduced set of features and determine comparative accuracy and computational time

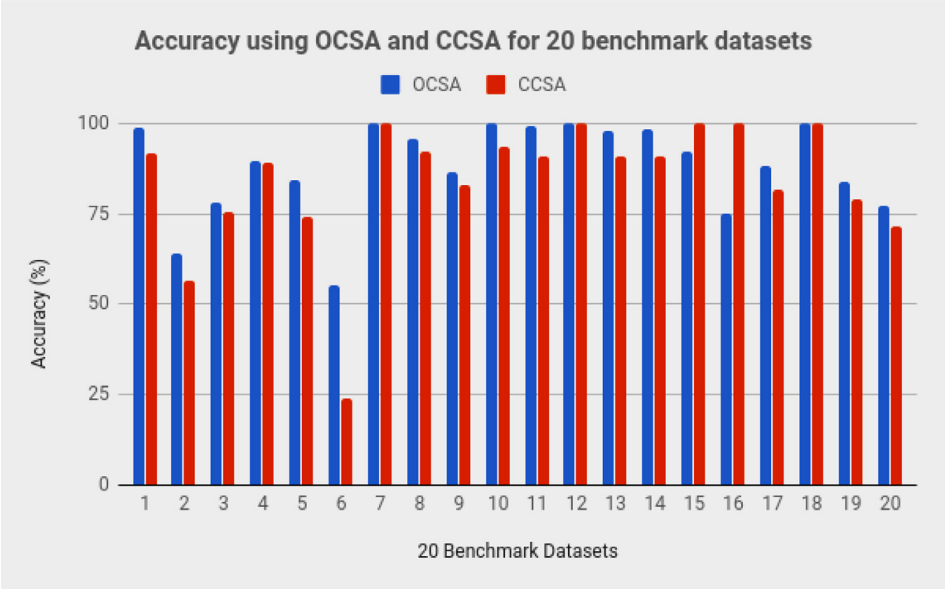


Fig. 8. Plot of the accuracy of 20 benchmark datasets for OCSA and CCSA.

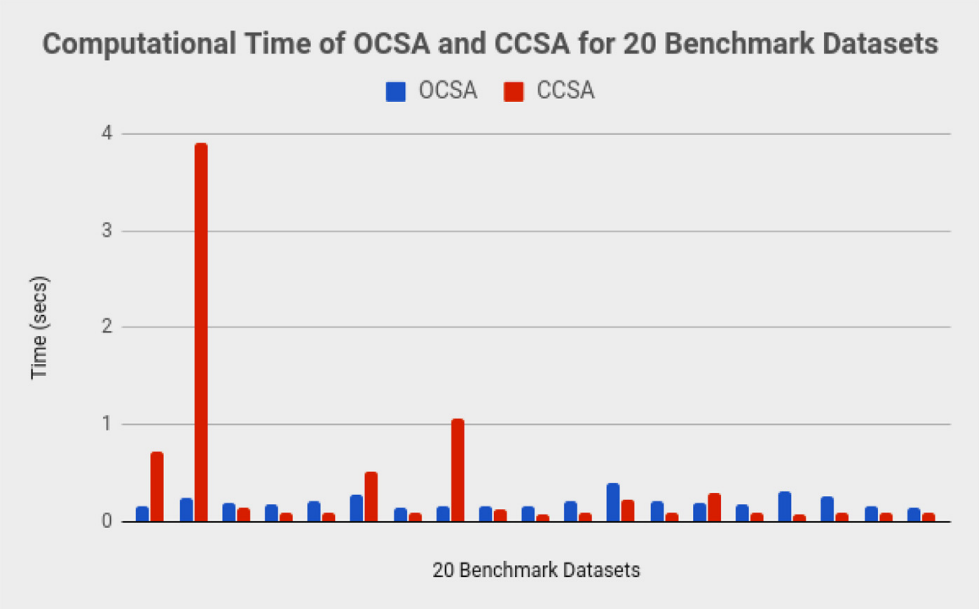


Fig. 9. Plot of the computational time of 20 benchmark datasets for OCSA and CCSA.

without degrading the performances of the system. The proposed algorithm is given as Algorithm 3 in Section 3. Diagnosis of Parkinson's disease at early stages is essential and thus to diagnose the disease, the proposed novel metaheuristic algorithm has been implemented and the results show that the proposed algorithm predicts the disease with an accuracy of 100%. Thus the method presented in this work helps individual to have proper treatment at an early stage. The Parkinson's disease and 20 benchmark datasets given in Tables 2 and 3 respectively have been applied to the proposed optimized crow search algorithm via K-Nearest Neighbor, Random Forest, and Decision Tree models to obtain the reduced set of features, accuracy and computational time. The novel optimized crow search optimization algorithm has been generating efficient and better results if it is used with Random Forest model. The performance comparison between the proposed algorithm and original chaotic crow search algorithm have been performed for 20 benchmark datasets and the results shows that the new proposed algorithm has selected fewer numbers of features with an accuracy of 88.2% and computational time of 0.20 s as compared to the accuracy of 84.2% and computational time of 0.40 s for the original chaotic crow search algorithm. The convergence rate also shows the stability of the fitness function of the proposed algorithm.

Algorithm 3
Optimized crow search algorithm.

Input: Different datasets with a different set of features associated with each crow and chaotic map as sine.

Output: Subset of feature and accuracy for the different machine learning models using an optimized version of binary crow search algorithm.

```
Set the value of m, Ap, fl, wf, tmax, d
Assign the dataset to the variable df
Split the dataset df into 75% - 25% ratio of training and testing data respectively
Assign the feature importance of each feature to the variable im
Initialize the position, xn, and memory, mem,  of the crows with the init function
for t in range( tmax ):
    Select  random crow num for chasing
        for i in range( m ):
            Get the random value between 0 and 1 and store it in r
            if r >= Ap then
                xnew.loc[ i , : ] ← xn.loc[ i , : ] + fl * r * (mem.loc[ num , :] - xn.loc[ i , : ] )
            else
                xnew.loc[ i,: ] ← a random number between 0 and 1.
    xnew ← 1 / ( 1 + exp( ( 10 * ( xnew - 0.5 ) ) ) )
        for i in range( m ):
        for j in range(  d ):
                if xnew[ i , j ] >= random() then
                    xnew [ i, j ] ← 1
                else
                    xnew [ i, j ] ← 0
    xn ← xnew
    for i in range( m ):
        Check if all position of a crow i is zero or not
        if  all position is zero ; replace every position of that crow with value 1
        ft_new ← fitness( xn )
    for  i in range( m ):
        if ft_new[ i ]>ft[ i ] then
                Update the memory : mem = xn
            for i in range( m ):
        Fit the model on the feature subset mem[ i, : ]
        Predict the model and then get the accuracy
Print the best accuracy with the number of features selected
```

The proposed algorithm can be applied in different relevant areas of research. We have used the proposed algorithm for diagnosis of Parkinson's disease based on handwriting exams. Research and practitioners can apply the proposed algorithm based on other types of diagnosis of PD. Researchers and practitioners can also optimize CCSA or other types of evolutionary algorithm for the diagnosis of Parkinson's disease.

References

[1] Pereira CR, Rosa GH, Weber SAT, Hook C, Papa JP. Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. 29th Conference on graphics, patterns and images (SIBGRAPI), 4–7 October 2016, São José dos Campos. Brazil: São Paulo; 2016.

[2] Pereira CR, Pereira DR, Silva FA, Masieiro JP, Weber SAT, Hook C, Papa JP. A new computer vision-based approach to aid the diagnosis of Parkinson's disease. Comput Methods Programs Biomed 2016;136:79–88.

[3] Afonso LCS, PEREIRA CR, Papa JP. Parkinson's disease identification through deep optimum-path forest. 30th conference on graphics, patterns and images (SIBGRAPI), 17–20 October 2017. Brazil: Niterói, Rio de Janeiro; 2017.

[4] Askarzadeh A. A novel metaheuristic method for solving constrained engineering optimization problem: crow search algorithm. Comput Struct 2016;169:1–12.

[5] Holland J. Adaptation in natural and artificial system. Ann Arbor: University of Michigan Press; 1975.

[6] Yang XS. A new metaheuristic bat-inspired algorithm. Trumpington Street, Cambridge CB2 1PZ, UK: Department of Engineering, University of Cambridge; 2010.
[7] Kennedy J, Eberhart RC. Particle swarm optimization. Proc of IEEE international conference on neural networks. 1995. p. 1942–8.
[8] https://en.wikipedia.org/wiki/Corvus_%28genus%29.
[9] P. Rincon, Science/nature|crows and jays top bird IQ scale, BBC News.
[10] Gonzalez-Pardo A, Jung JJ, Camacho D. ACO-based clustering for ego network analysis. Future Gener Comput Syst 2017;66:160–70.
[11] Ramirez-Atencia C, Mostaghim S, Camacho D. "A knee point based evolutionary multi-objective optimization for mission planning problems. Proceedings of the genetic and evolutionary computation conference. 2017. p. 1216–23.
[12] Ramirez-Atencia C, Bello-Orgaz G, Moreno DR, Camacho D. Solving complex multi-UAV mission planning problems using multi-objective genetic algorithms. Soft Comput 2017;21(17):4883–900.
[13] Martin A, Menéndez HD, Camacho D. Genetic boosting classification for malware detection. Evolutionary computation (CEC), 2016 IEEE Congress on. 2016. p. 1030–7.
[14] Martín A, Fuentes-Hurtado F, Naranjo V, Camacho D. Evolving deep neural networks architectures for android malware classification. Evolutionary computation (CEC), 2017 IEEE congress on. 2017. p. 1659–66.
[15] Martín A, Menéndez HD, Camacho D. MOCDroid: multi-objective evolutionary classifier for android malware detection. Soft Comput 2017;21(24):7405–15.
[16] Gonzalez-Pardo A, Palero F, Camacho D. An empirical study on collective intelligence algorithms for video games problem-solving. Comput Inf 2015;34:233–53.
[17] González-Pardo A, Palero F, Camacho D. Micro and macro lemmings simulations based on ants colonies. European conference on the applications of evolutionary computation. 2014. p. 337–48.
[18] Fierrez J, Morales A, Vera-Rodriguez R, Camacho D. Multiple classifiers in biometrics. Part 1: fundamentals and review. Inf Fusion 2018;44:57–64.
[19] Fierrez J, Morales A, Vera-Rodriguez R, Camacho D. Multiple classifiers in biometrics. Part 2: trends and challenges. Inf Fusion 2018;44:103–12.
[20] Črepinšek M, Liu SH, Mernik M. Exploration and exploitation in evolutionary algorithms: a survey. ACM Comput Surv 2013;45. http://dx.doi.org/10.1145/2480741.2480752. Article 35.
[21] Nakamura RYM, Pereira L, Costa K, Yang XS. BBA: a binary bat algorithm for feature selection. Brazil: Department of Computing Sao Paulo State University Bauru; 2012.
[22] Gupta D, Ahlawat AK. Usability feature selection via MBBAT: a novel approach. J Comput Sci 2017https://doi.org/10.1016/j.jocs.2017.06.005.
[23] Sayed GL, Hassanien AE, Azar AT. Feature selection via a novel chaotic crow search algorithm. Neural Comput Applic 2017https://doi.org/10.1007/s00521-017-2988-6.
[24] Prior H, Schwarz A, Güntürkün O. Mirror-induced behavior in the magpie (pica pica): evidence of self-recognition. PLoS Biol 2008;6(8):e202.
[25] Bache K, Lichman M. UCI machine learning repository Retrieved July 19 http://archive.ics.uci.edu/ml; 2016. Retrieved July 19.

Deepak Gupta has completed B.Tech. from Guru Gobind Singh Indraprastha University, M.E. from Delhi College of Engineering and Ph.D. from Dr. APJ Abdul Kalam Technical University. He has published 25 text books on various topics of computer science & engineering with International Publishers, has edited two books published by Elsevier and 21 research papers in various International Journals of repute.

Shirsh Sundaram is an eminent student of Computer Science department at Maharaja Agrasen Institute of technology, Delhi. He is a passionate and dedicated Machine Learning enthusiast and developer who loves to make impact to the society. He is also looking forward for some more challenging tasks.

Ashish Khanna is a highly qualified with around 15-years of rich expertise in Teaching and Research & Development with specialization in CSE Subjects. He received his Ph.D. from NIT, Kurukshetra. He has completed his M.Tech. and B.Tech. from GGSIPU. He has published many research papers in reputed journals. He also has papers in SCI indexed and Springer journals.

Aboul Ella Hassanien is Professor at Cairo University and the chair of the technology center of blind and visual impaired people. Abo received his B.Sc. with honours and M.Sc. degree, both from Ain Shams University. On September 1998, he received his doctoral degree from the Department of Computer Science, Graduate School of Science & Engineering, Tokyo Institute of Technology, Japan.

Victor Hugo C. de Albuquerque has a Ph.D. in Mechanical Engineering with emphasis on Materials from the Federal University of Paraíba, an M.Sc. in Teleinformatics Engineering from the Federal University of Ceará, and he graduated in Mechatronics Technology at the Federal Center of Technological Education of Ceará. He is currently Assistant VI Professor of the Graduate Program in Applied Informatics.

Periodicals of Engineering and Natural Sciences ISSN 2303-4521
Vol.6, No.1, May 2018, pp. 107~119
Available online at: http://pen.ius.edu.ba

Mechanism of Joint Formation and Characteristics of Interface in Ultrasonic welding: Literature Review

Bharat Sanga[1], Reeta Wattal[2], D. S. Nagesh[3]
[1] Faculty, Department of Mechanical Engineering, Guru Nanak Dev Institute of Technology, Delhi, India
[2,3] Professor, Department of Mechanical Engineering, Delhi Technological University, Delhi, India

Article Info

Article history:

Received Jan 27, 2018
Revised May 3, 2018
Accepted May 14, 2018

Keyword:

Ultrasonic welding
Ultrasonic metal welding
Intermetallic compound
Heat affected zone
Thermo-mechanical affected zone
Welding interface

ABSTRACT

The study of ultrasonic welding has been going on for more than 50 years. The mechanism of joint formation and characterization of the interface in ultrasonically made joints between metal sheets and wires have been the most thought-provoking area for the researchers. The study of post-welding microstructure, the formation of any intermetallic compound at the interface and their effect on the joint strength, the presence of heat affected zone in the ultrasonically joined sheets has been explored but still, arguably the least understood. Interface characteristics are different in similar and dissimilar combinations of metals. This work presents a comprehensive review of literature regarding the studies on the microstructural analysis at the interface of the joints made by ultrasonic welding on different combinations of metal sheets. Additionally, this paper provides an analysis of the observations made by different scientists that promotes the future scope of research in this area. The study has been confined to ultrasonic metal welding only.

Corresponding Author:

Bharat Sanga[1],
Department of Mechanical Engineering,
Guru Nanak Dev Institute of Technology,
Sector-15, Rohini, Delhi-110085, India.
Email: bharatsanga_2k14@dtu.ac.in

1. Introduction

Joining and adhesion of parts are essential steps in the assembly of components. Joining of different materials has been a challenging task for engineers for a long time. Many applications in the field of automobiles, aeronautics, electronics, shipbuilding and architecture require joining of dissimilar materials such as metal to non-metal [1]. Miniaturization and weight reduction is a major requirement of today's product development phases. The need is fulfilled by unconventional and newer joining techniques. One such joining technique gaining popularity, particularly in the automotive industry is Ultrasonic Welding (USW) [2].

Introduced in the early 1950s, industrial use of USW has started in the year 1963. It was initially used for wire bonding and tube sealing, but the technique has gained popularity due to its unique advantages such as joining between dissimilar materials, power saving, extremely low welding time and capability to join very thin sheets using lighter and more versatile equipment to produce a stronger, smaller and precise weld [3,4].

In this process, the material to be joined is held together under moderate pressure and a high-frequency ultrasonic vibration is applied locally on the joint. Due to the combined effect of the pressure and ultrasonic vibrations, the surface asperities, contaminants and oxides between the overlapping surfaces get removed and

the pure metallic surfaces come into contact with each other, resulting in a sound metallurgical bond. The USW is used both for joining of plastic parts as well as metal sheets, popularly known as Ultrasonic Metal Welding (USMW) [5].

Researchers have been investigating different characteristics of USW, but the activities occurring at the joint interface have not been fully explored yet. In this paper, a comprehensive literature review of the joining mechanism taking place in case of USMW has been presented. The selection of study has been focused only on ultrasonic spot or seam welded joints of similar or dissimilar combinations of thin metal sheets. Ultrasonic consolidation or ultrasonic additive manufacturing, ultrasonic wire bonding, ultrasonic plastic welding and ultrasonic hybrid joints (metal to non-metal combinations) are not included in this study.

2. Theoretical Background

USW equipment has four main parts (Fig. 1): A Power Supply converts low-frequency electricity (50-60 Hz) to high-frequency electricity (20 - 40 kHz). The transducer or Signal Converter changes the high-frequency electricity into high-frequency ultrasound waves. A Booster increases the amplitude of the vibrations. Finally, a Horn or **Sonotrode** focuses the ultrasonic vibrations and delivers them to the materials to be welded. The horn is tuned to operate in the longitudinal mode and its precise design and characterization are required to ensure the required dynamic characteristics of the device. Besides these parts, there is an anvil upon which the materials to be welded are placed and held together during welding with the help of air pressure supplied by a pneumatic system.

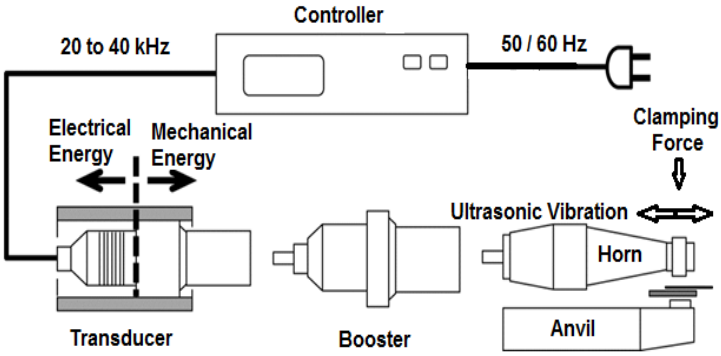


Figure 1: Ultrasonic Metal Welding Set-up

A number of parameters can affect the welding process, such as ultrasonic frequency, vibration amplitude, clamping force, power, energy, welding time, materials, part geometry, and tooling [1,4]. Most USMW systems operate at 20 to 40 kHz, with 20 kHz being the most common frequency since, at this frequency, a very high strain rate is achieved in a fraction of a second and the micro-sized interfacial surface asperities are removed instantaneously [5].

3. Microstructural Analysis

Microstructural analysis has been the area of interest for the scientists right from the early days of the invention in the field of USW. The works performed by Okada et al. [6], Daniels [7] and Hazlett et al. [8] included fundamental studies on joining of similar and dissimilar combinations of pure metals such as Al, Cu, Zn, Ti, Be, Ag and Ni and alloys like Brass and SS. The sheet thickness had been in the range of 0.06-0.5 mm.

A common observation made by all the above scientists was that the temperature of the interface has reached above the recrystallization temperature and a recrystallization texture was observed but no Inter-Metallic Compound (IMC) or other phases were found to be present at the interface (Fig. 2). However, no common observation was found in regard of inter-atomic diffusion. Inter-atomic diffusion was observed between Cu and Ti by Okada [6], but the technique used was X-ray micro-analysis. Later, Daniels [7] confirmed that no

melting or diffusion was observed in Cu to Ni welding. Hazlett's [8] observations based on Scanning Electron Microscope (SEM) images supported the observations of Daniels [7] to some extent. He concluded that "diffusion" supposed to be the main joining mechanism at the interface was due to grain boundary rather than bulk diffusion. Mechanical mixing at the interface and the accompanying metal bonding at the nascent interface were found to be the main causes of joint formation.

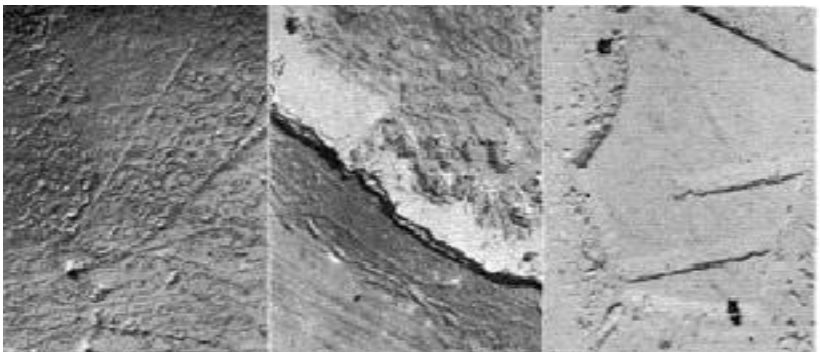


Figure 2: Microstructure of ultrasonically welded 0.5 mm Cu and 0.5 mm Ti showing no IMC or other phases [6]

The findings of H. Kreye [9] were different from that of above researchers. He investigated the joining phenomenon of Cu and Al alloy sheets joined by different solid state welding techniques viz. Ultrasonic (sheet thickness: 0.4 mm), explosion and friction welding. He found that microstructural changes have occurred during welding, resulting in the formation of new grains along the interface as shown in Figure 3.

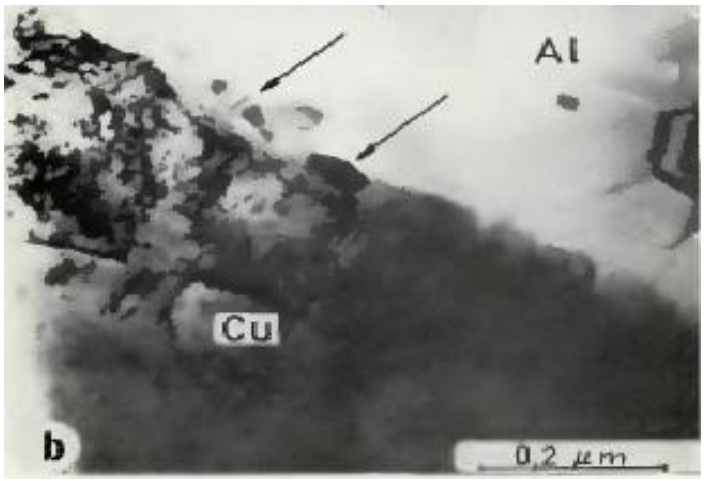


Figure 3: TEM micrograph of US welded Cu /Al interface. Arrows show formation of new grains [9]

It was observed during ultrasonic welding that bonding occurred predominantly by melting of a very thin and narrow surface layer (thickness < 1 µm) for a very short time. The small grain size of the particles and Co particles taking more time in dissolution supported this theory. Co particles take more time in dissolution when the temperature of Co reaches above its melting temperature. No IMC was observed due to rapid heat transfer from the bonding area to the surrounding atmosphere. The melting phenomenon was also observed by Y. C. Chen et al. [12] and Bakavos and Pragnell [13].

However, H. Kreye's observation contradicts the joining mechanism explained by many other researchers. Al-Sarraf Z. et al. [10] performed USW with different combinations of 0.1 mm and 0.5 mm sheets of Al and Cu and found that there was diffusion of Al on Cu side. SEM images revealed that more Al deposited on the Cu side as compared to less deposition of Cu on Al side due to Cu being a softer metal of the two. But

definitely, there was an interaction at the atomic level between the atoms of interacting surfaces during ultrasonic welding. This fact was proved by Watanabe T. et al [11]. He used Auger Electron Spectroscopy (AES) to analyze the interface of ultrasonically joined 0.5 mm thick Al sheet with Alumina (Al_2O_3) and observed the formation of a chemical bond between the atoms of Al and Oxygen of Al_2O_3.

Y. C. Chen et al. [12] and Bakavos and Pragnell [13] studied the interfacial microstructural development in ultrasonic spot welding of 0.93 mm thick AA6111-T4 automotive sheet and contradicted the general concept that no Heat Affected Zone (HAZ) is present in the USW owing to low-energy input. They have described the term 'Thermo-mechanically Affected Zone' (TMAZ) for the area covered between the sonotrode tip and anvil and 'HAZ' for the area outside the anvil. The same concept was also given by M. Shakil et al. [14]. TMAZ is shown in Figure 4.

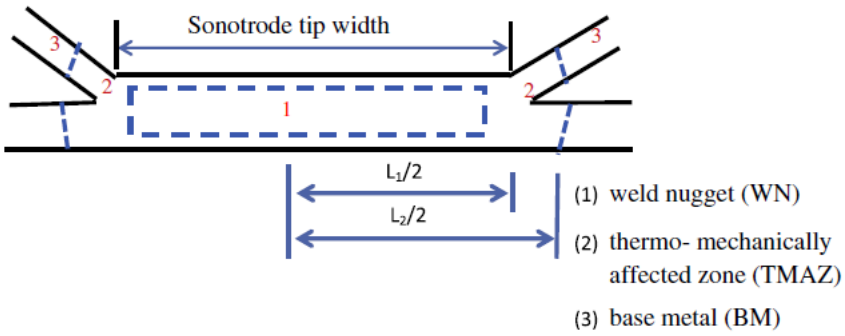


Figure 4: TMAZ was shown [14]

The hardness was low in the centre of the weld area and increased as moved away from the centre. The hardness of the samples was recovered to their normal T4 level after 2 weeks of natural aging. A fine grain boundary precipitation was found when the maximum rise in the temperature was in the range of 250 – 400 °C. But at a higher temperature level, such as above 500 °C, there was no grain boundary precipitation due to melting and mixing of the parent materials as shown in Figure 5.

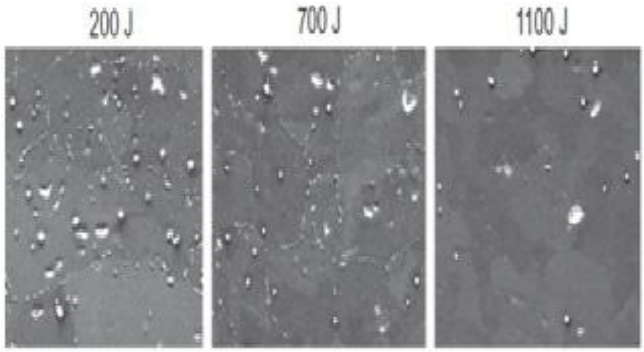


Figure 5: SEM images of grain boundary precipitation in welds produced with increasing energy showing fine grain boundary precipitation [13]

Another aspect observed during this research was that natural aging in AA6111 alloy remained continued even after USW as it would be in normal circumstances for conventional solution treated sheets. This phenomenon was termed as "Accelerated Natural Aging". This concept is also supported by Haddadi [23].

The initiation of the bonding started with the formation of micro-bonds at the points where the oxide was broken or the galling of the asperities occurred. The micro-bond density increases with more energy input and they coalesce to form more bonds. During this initial phase, the bonds break and reform again and the flakes of oxides were also included at the interface. At higher energy level, the swirl-like interface surface gets folded due to the wavelike displacement of intercalated materials (Fig. 6).

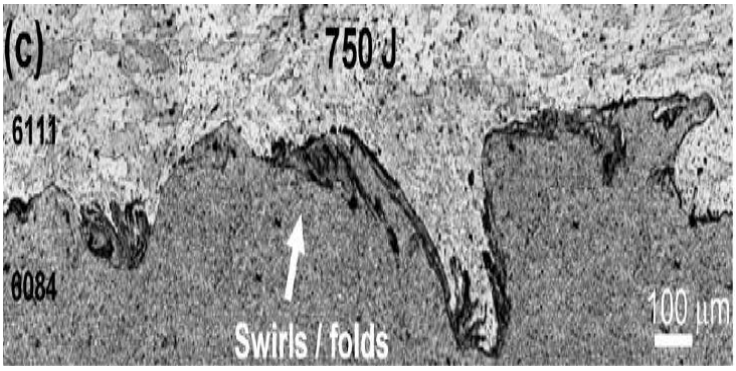


Figure 6: Enlarged region from weld cross section showing interface Swirls/folds [13]

The work of M. Shakil et al. [14] involves USW of 300 µm thick sheets of AA3003 alloy and 50 µm thick foil of 304 SS alloy using 2.4 kW ultrasonic spot welding machine. The optical micrographs showed convoluted waves like appearance of weld interface and local micro-welds (Fig. 7) similar to the observations made by Y.C.Chen et al. [12] and Bakavos and Pragnell [13].

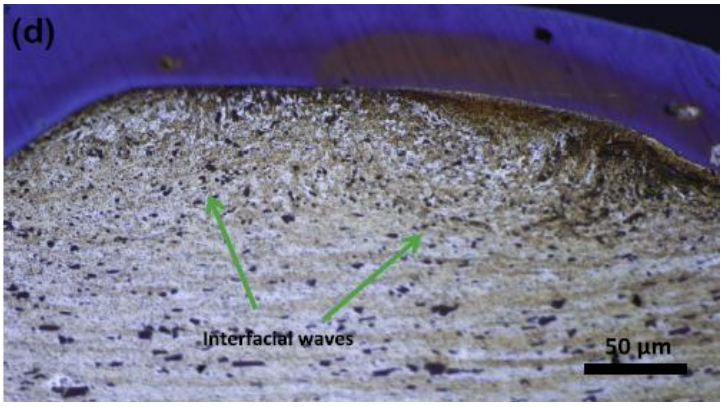


Figure 7: Optical images showing interfacial waves and micro welds [14]

Bonding mechanism was summarized in three steps as (i) the formation and expansion of the bonding line along the weld interface (ii) deformation along the bonding line in TMAZ and (iii) micro-bonding growth between the materials. It was also observed that the elongated grain microstructures disappear, and recrystallization takes place by plastic deformation. Hence, it was concluded that the material at the interface receives more ultrasonic energy and its temperature rises continuously due to which instant work hardening and softening take place simultaneously. Thus, it was compared with the cold work followed by annealing process.

Recovery and recrystallization were also observed by J. Magin et al. [15] during the ultrasonic torsion welding of 0.8 mm thick pure aluminum sheet AA1199 and commercially pure Titanium grade 1 Ti sheet. It was confirmed that melting did not take place and the joint was formed well below the melting point in solid state without the formation of any IMCs. Similar results were also derived by Zhang C. and Li L. [16] during USW of Al foils. They simulated the coupled thermo-mechanical fields using a 3-D Finite Element Analysis (FEA) model to study the effect of heat generation and plastic deformation at the interface and concluded that it is only the severe, localized plastic deformation which is responsible for the bond formation.

The work of C.Q. Zhang et al. [17,18] explored USW of 0.93 mm thick Al alloy AA6111-T4 and 1.0 mm thick Ti alloy $TiAl_6V_4$ in 2014 and again 1.0 mm thick Al alloy AA2139-T8 and 1.0 mm thick Ti alloy $TiAl_6V_4$ in 2016. Almost similar results were obtained. No visible IMC layer was detected. It was observed that Al grains near the weld interface have been distinctly refined by the ultrasonic spot welding process, but

there was no difference in the grains of $TiAl_6V_4$. Grain refinement was also observed by Prangnell et al. [19] during USW of 1.0 mm thick Al 6111 sheet with automotive DC04 steel sheet. However, Zhang suggested that little deformation occurred on the Ti side due to its high strength at peak temperature. The absence of IMC in Al-Ti combination was attributed to three factors: (i) very less solid solubility (ii) relatively high activation energy and (iii) the higher energy barrier to nucleation of the IMC phase. In a similar work done by Zhu Z. et al. [20] on AA6061 and $TiAl_6V_4$, it was observed that diffusion across the interface took place and thickness of diffusion layer increased with increasing welding time. No description of IMCs formed during the process was given. Frank Balle et al. [21] also ultrasonically welded 1.0 mm thick Al alloy AA7075 and $TiAl_6V_4$ and observed formation of micro-welds which spread along the joint interface leading to the solid state joint formation. However, they did not investigate any IMC or reaction layer in their work.

The findings of S.Q. Wang et al. [22] were contradictory to that of C.Q. Zhang et al. [17,18]. The ultrasonic joining of 1.5 mm thick AA 5754 alloy and 1.0 mm thick $TiAl_6V_4$ alloy with an interlayer of 75 µm thick Al was performed. SEM and X-ray Diffraction (XRD) analysis of the microstructure revealed that IMC was formed, but out of many possible combinations, only $TiAl_3$ IMC phase was present due to its ease of formation at a comparatively low temperature as shown in Figure 8. The joint mechanism was attributed to the melting of the Al interlayer and Al alloy. However, Ti alloy, being harder than Al, was not found in the molten form, but got dissolved in the liquefied Al at that temperature.

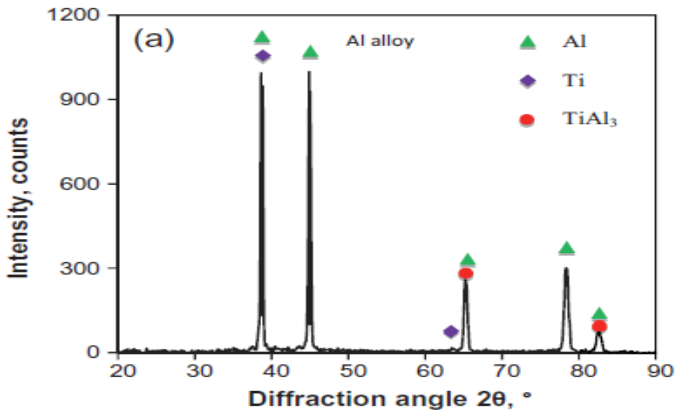


Figure 8: XRD pattern of ultrasonic welded Al/Ti alloy joint with Al interlayer [22]

Ultrasonic weldability of Al alloy AA6111-T4 with bare and zinc coated Steel of approximately 1 mm thickness was investigated by Haddadi and Abu-Farha [23]. Their investigation in reference of IMCs, also supported the findings of S.Q. Wang et al. [22]. The growth of Al-Fe intermetallic layer was seen spreading discontinuously along the interface. The presence of IMC was verified by the Energy Dispersive Spectroscopy (EDS) analysis. Several Al-Fe IMCs such as $FeAl_3$, Fe_2Al_5 and FeAl were observed, Fe_2Al_5 having the largest proportion. The formation of the IMCs was co-related to the fracture of the joint in lap-shear test. At sufficiently high weld time, the failure occurred by cracking through the interface of the reaction layer and the steel sheet where brittle FeAl phase was formed. A significant HAZ/TMAZ was also observed like Y.C. Chen [12], D. Bakavos and P.B. Pregnell [13] and M. Shakil [14].

Shin-ichi Matsuoka et al. [24] joined same and different combinations of Al alloys (AA1050H-24, AA1050-O and AA3003-O) and Cu alloy (C1020) of thickness ranging from 0.3 mm to 1.5 mm. The SEM observation of weld between AH/A1-O showed that the welding interface had no defects such as cracks or exfoliation. The presence of oxides or organic films was observed in incomplete welding regions, but the interface of good weld was free from such defects. The surface in the vicinity of the joint appeared as pleat-like traces due to pressure and frictional heat of the horn. Energy Dispersive X-ray (EDX-ray) analysis observed the presence of a transit layer of approximately 2 µm thickness on both sides of the interface which indicates the possibility of formation of some reaction or diffusion layer (Fig.9). Formation of IMCs or other welding attributes such as HAZ/TMAZ were not explored.

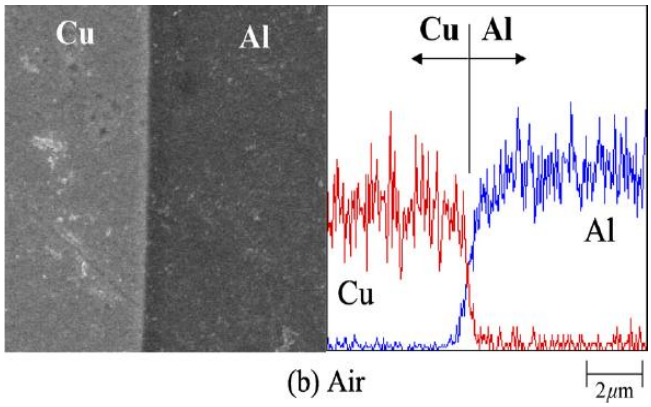


Figure 9: SEM image showing formation of 2 μ thick transit layer of both alloys (Al/Cu) [24]

Fujii [25] ultrasonically joined 1.0 mm thick sheets of AA1050 Al alloy and Cu. He strongly advocated the recrystallization of grains at the interface. Recrystallization was attributed to the creation of large dislocation density due to deformation of the grains. The thickness of the reaction layer of Al_2Cu IMC phase was found to be in the range of 40-100 nm and it helped in increasing the joint strength, unlike normal IMC, due to its very less thickness (< 1μm). The same observations were made by the same author [26] during the USW of 1.0 mm thick AA6061 Al alloy and 0.5 mm thick 304 SS sheets where a number of nanometer-size Fe-Al IMC formed discontinuously around the interface and contributed to the weld strength. The formation of Fe-Al IMC was attributed to pipe diffusion of Fe in Al around the interface due to excessive dislocation density at the interface owing to severe shear deformation during USW (Fig. 10). The effects of excessive dislocation density were explained in detail by Gunduz et al. [38].

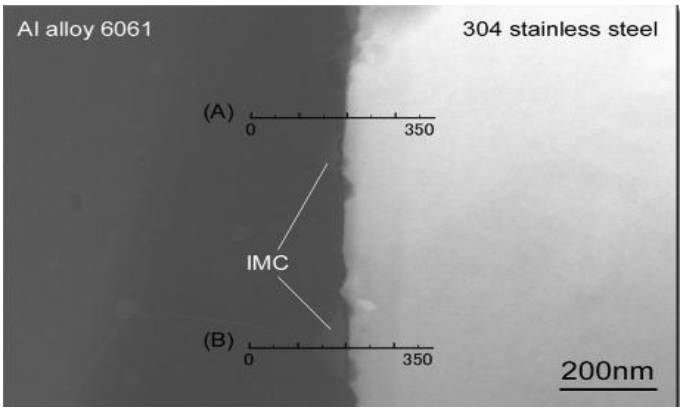


Figure 10: TEM image around the weld interface showing a thin layer of IMC [26]

During the similar experiment done by Y. Y. Zhao et al. [27] the ultrasonic joint was made between 1.0 mm thick sheet of AA 6061 and pure Cu. The joint formation was attributed to micro-welds at low energy input, but at higher energy input, deformation along with mechanical interlocking due to formation of swirl like pattern along the interface was found to be the main cause of joint formation. The joint strength increased with increasing welding energy but decreased at very high levels of welding energy. Similarly, no IMCs were observed at lower values of welding energy, but a very thin dark layer of Al_4Cu_9 IMC was found at very high energy level.

C.Y. Zhang et al. [28] and Xin Wu et al. [29] observed similar characteristics of interface during USW. C.Y. Zhang et al. [28] while doing the ultrasonic spot welding on a 2-mm thick wrought AA5754-O Al alloy sheet revealed through SEM images that as the energy input was increased; the waviness, branching, bifurcation and pores at the interface decreased and the strength of the joints increased. Xin Wu et al. [29] also observed the similar results during the USW of battery tabs made of Al and Cu. Highly wavy, curly and intermixed layers of Al were observed. However, the observations of R. Jahn [30] contrasted with the above and showed

that waviness, branching and bifurcation at the interface increased with higher energy during the USW of 0.9 mm thick Al alloy AA6111-T4.

R. Balasundaram et al. [31] conducted experiments on Al alloy and Cu alloy, with and without Zn inter-layer. 1.5 mm thick sheet of AA5754-O and 1 mm thick sheet of C110 Cu alloy and 50 µm thick pure Zn interlayer were used for experimentation. The SEM micrographs revealed swirling interface with the void when sheets were joined without interlayer, however, a flat and tight interface was obtained with Zn inter-layer. Secondly, no IMCs were produced without inserting Zn inter-layer, but a discontinuous IMC layer of Al_2Cu was observed mostly at Cu side due to higher solubility of Al in Cu when welded with Zn inter-layer (Fig. 11). Another aspect observed was the possibility of Liquid Metal Embrittlement (LME) where liquid Zn could embrittle Cu and a crack could be induced.

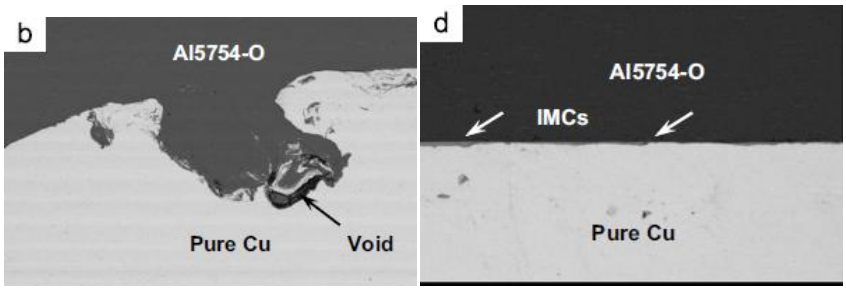


Figure 11: Micrographs showing Swirling interface in (b) and flat interface in (d). IMCs in Al alloy/Cu joints with Zn inter-layer shown in (d) [31]

The interface characteristics were also explored during the USW performed between different metals combined with Mg alloys. Formation of brittle IMCs was explored by A. Panteli et al [32] during the ultrasonic joining of 1.0 mm thick sheets of Al alloy AA6111 and Mg alloy AZ31-H24. He clearly observed thick layers of IMCs Al_3Mg_2 and $Al_{12}Mg_{17}$. The thickness of the IMCs increased with increasing welding energy. The rapid growth rate of IMC layers was attributed to the enhanced diffusion due to micro-cracking in the brittle reaction layer. They also observed the melting at the interface between the IMC phase and the Mg alloy base at higher energy levels and at decreased melting temperature.

V. K. Patel et al. [32,33,34] conducted many experimentations to study the weldability aspect of ultrasonically welded Mg alloy AZ31-H24 in similar and dissimilar combinations with other metals and with or without using interlayer. During the USW of 2 mm thick sheet of Mg alloy AZ31-H24 at different energy levels, he observed the complete recrystallization of microstructure and the grain size increased with respect to the increase in energy level (Fig. 12).

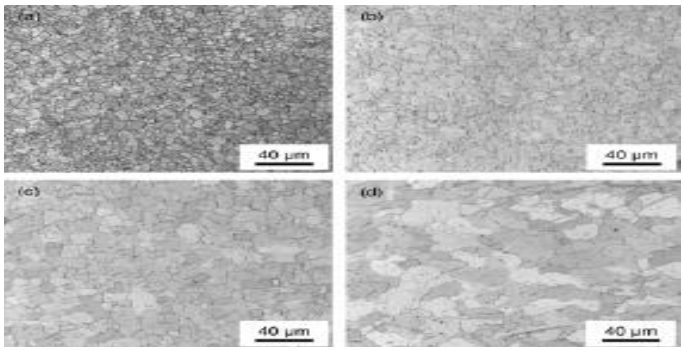


Figure 12: Microstructure of AZ31 at different energy levels compared to the base metal [32]

In another study [33] on the USW of 2 mm thick sheet of AZ31B-H24 Mg alloy and 0.8 mm thick sheet of High Strength Low Alloy (HSLA) with the 70 µm thick pure zinc sheet placed in-between as an interlayer, he observed very thin IMCs such as Mg_2Zn_{11}, Mg_7Zn_3, MgZn and $MgZn_2$ between Mg and Zn substrates but no IMCs were observed between Zn and Fe. Also, no oxide was found at the Mg / steel interface (Fig. 13).

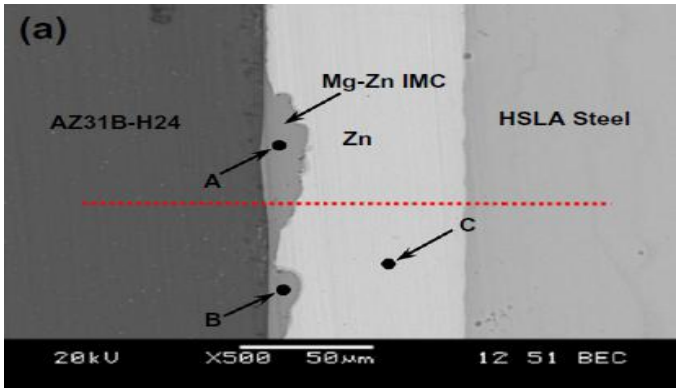


Figure 13: Microstructure of ultrasonic welded Mg/HSLA with Zn inter-layer showing IMCs between Zn and Mg alloy, but no IMCs on HSLA side [33]

So it was suggested that the joint at the steel side was formed by solid solution of Zn and Fe. However, during the similar studies conducted by the same researcher earlier (2012) with 2 mm thick AZ31B-H24, 1.5 mm thick AA5754-O and 50 µm thick pure Sn interlayer [34], it was observed that IMCs form at interface without using the interlayer but no brittle IMCs formed when Sn interlayer was placed in between. This improvement was attributed to the formation of solid solutions of Sn with Mg and Al and due to formation of eutectic structure such as Mg_2Sn in the interlayer (Fig. 14). The joint strength was improved as observed by Fujii [24].

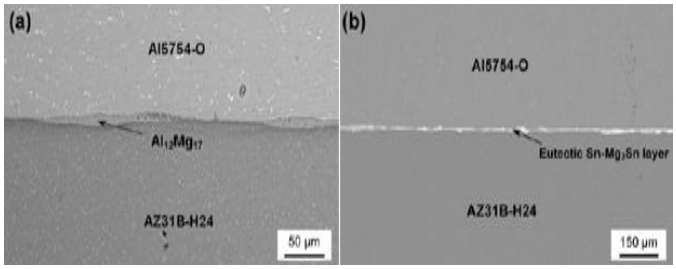


Figure 14:(a) Micrograph showing interface with IMCs without using interlayer (b) formation of solid solution of Sn with Mg and Al at the interface [34]

Similar observations were made by Watanabe et al. [35] while making USW joints with 0.8 mm thick SS400 mild steel sheet and 1.2 mm thick AA5052-H24 Al alloy sheet containing Mg with and without using commercially pure AA1050 alloy having 1.2 mm thickness as an insert material. It was observed that the strength decreased due to the formation of IMC Fe_2Al_5 at the interface when welded at high welding time without using insert material. Use of insert material improved the joint strength significantly since no IMCs were formed even at a higher welding time between the interfaces of SS400/AA1050 and AA1050/AA5052 (Fig. 15).

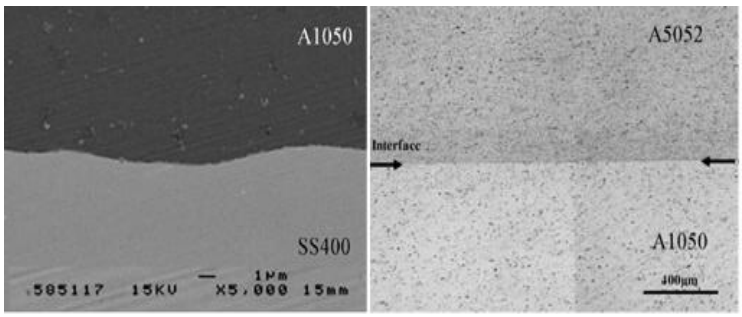


Figure 15: SEM image showing no IMCs produced at the interfaces with insert material [35]

Similar investigations were made on interface microstructure of Mg alloy/Cu combination by Macwan and Chen [36]. Dissimilar pairs of 2 mm thick AZ31B-H24 and 1 mm thick pure Cu were ultrasonically welded. Continuously increasing diffusion layer of IMCs at the interface was observed on the Mg side, but not on the Cu side, due to higher melting temperature and thermal conductivity of Cu. Interface thickness increased with increasing temperature (Fig. 16).

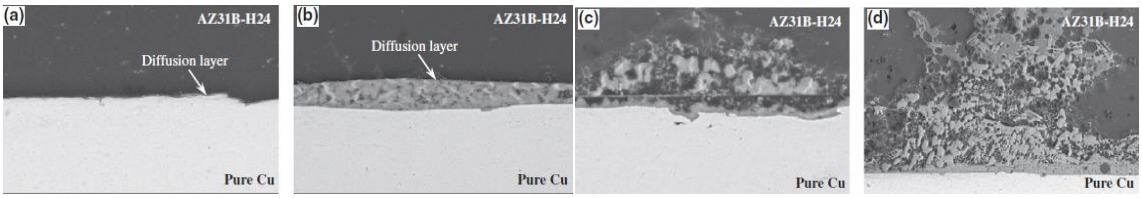


Figure 16: Micrographs showing increasing diffusion thickness at the interface [36]

The interface consisted of eutectic structure of Mg and Mg_2Cu. The thickness of diffusion layer was observed significantly larger as compared to other instances. This enhanced dynamic diffusion was attributed to the diffusion-induced vacancies, the phenomena, also observed by I. E. Gunduz et al [37] during the US welding of 50 µm thick 1100-H19 Al foil and 250 µm thick 99.8% pure Zinc sheet. Fujii et al [24] also observed the same phenomenon in their work. A large reduction in the melting point was observed due to strain-induced excess vacancy that resulted in local melting at the interface even at lower temperature and short welding time in USW.

M. P. Satpathy et al. [38] explored the welding interface during the USW of 0.3 mm thick sheets of Al alloy AA1100 and UNS C27000 Brass and categorized the joint as "under", "good" and "over" on the basis of the condition of the interface. A good weld had high density of micro-bonds and was free from gaps between sheets. Inter-atomic diffusion was observed as the main reason for joint formation.

4. Observations

This review paper critically reviews the observations made by researchers regarding the joining mechanism and the characteristics of the joint interface produced by ultrasonic metal welding of similar as well as dissimilar metal combinations. The following observations are made:

1. The bonding mechanism in USW has been the core interest for many researchers, but still, no common consensus on this issue has been made.

2. Following bonding mechanisms were observed- Interatomic diffusion [5,7,9,19,23,31,32,33,34,36,37,38], recrystallization and recovery [13,14,24,32,33], melting at the interface [8,12,21,22,30,31] mechanical mixing/interlocking at interface [7, 26], and micro-welding [11,13,20,26].

3. No IMCs were observed in the combinations of pure metal such as Cu, Ti, Ag, Be, Al, Ni, Zn [5,6,7,8,9].

4. IMCs were observed in the combinations of AA6111 and Zn coated Steel [22], AA5754 and $TiAl_6V_4$ [21], AZ31B and HSLA [33] and AA5052 and SS400 [35], AA6061 and SS304 [25], AA6061 and Pure Cu [26], Al alloy and Cu alloy with Zn as inter-layer [30], Al alloy AA6111 and Mg alloy AZ31-H24 [31].

5. Formation of thick layers of brittle IMCs can be avoided by using some suitable interlayer [34]. Nevertheless, thin layer of IMCs contributed to the strength of the joint [24, 25].

6. Few researchers observed HAZ and TMAZ [11, 13, 22] contradicting the conventional concept of the absence of HAZ / TMAZ in solid state welding.

7. Joint failures were attributed to the presence of IMCs [21, 22, 25].

8. Enhanced diffusion was observed due to diffusion induced vacancies [24, 25, 31, 37].

9. Characterization of the interface as described by many researchers is – convoluted wavelike or swirling, branched, bifurcated and free from any cracks, exfoliation or oxides [13, 23, 27, 28, 29, 30].

5. Conclusions

All the studies have shown that there are many mechanisms working behind the formation of the ultrasonically welded joint. In some combinations, there are more than one joining mechanisms taking place, whereas in other cases, there is only one. Secondly, different metal combinations show the different possibilities of joining mechanisms. Still, there is ample scope for finding some uniform theory which can predict the joint mechanism.

It can be concluded that IMCs are formed in the dissimilar combinations of metals or alloys which reduce the joint strength in most of the cases but a suitably chosen inter-layer may improve the joint strength in some of the metal combinations by making a very thin layer of IMCs of the parent metals.

A unique concept of enhanced dynamic diffusion due to accelerated dislocation assimilation has been found to be the main cause of severe deformation taking place during ultrasonic welding. This phenomenon reduces the melting point of the parent metals and promotes the ‘welding due to melting’ concept contradicting the ‘no melting in cold welding’. Similarly, the presence of HAZ and TMAZ indicates that ultrasonic metal welding may be assumed to be an extension of fusion welding and opens the need for further exploration.

References:

[1] Elma Kicukov and Ali Gursel, “Ultrasonic welding of dissimilar materials: A review,” *Periodicals of Engineering and Natural Sciences*, Vol. 3, No. 1, pp. 28-36, 2015.

[2] K. Graff, Ultrasonic Metal Welding, *New Developments in Advanced Welding,* Cambridge, pp. 241-269, 1997.

[3] American Welding Society Handbook, Vol. 2, 8th ed., 1997.

[4] C. F. Brockelsby, “Some Applications of Ultrasonics,” *J. Sci. Instrum.*, 40, pp. 153-157, 1963.

[5] E. De Vries, *Mechanics and Mechanisms of Ultrasonic Metal Welding,* Ph.D. Thesis, Ohio State University, 2004.

[6] Minor Okada, S. Shin, M. Miyagi and H. Matsuda, “Joint mechanism of ultrasonic welding,” *Journal of Japan Institute of Metals,* 4(26), pp. 250-256, 1962.

[7] H. P. C. Daniels, “Ultrasonic Welding,” *Ultrasonics*, pp. 190-196, 1965.

[8] T. H. Hazlett and S. M. Ambekar, “Additional Studies on Interface Temperatures and Bonding Mechanisms of Ultrasonic Welds,” *Welding Research Supplement, AWS,* pp. 196s-200s, 1970.

[9] H. Kreye, “Melting Phenomena in Solid State Welding Processes,” *Welding Research Supplement, AWS,* pp. 154-s to 158-s, 1977.

[10] Z. Al-Sarraf and M. Lucas, “A study of weld quality in ultrasonic spot welding of similar and dissimilar metals,” *Modern Practice in Stress and Vibration Analysis, Journal of Physics: Conference Series,* 382, pp. 1-6, 2012.

[11] T. Watanabe, A. Yanagisawa and S. Sunaga, "Auger Electron Spectroscopy Analysis at the Ultrasonically Welded Interface between Alumina and Aluminum," *Metallurgical and Materials Transactions A*, 34(A), pp. 1107-1111, 2003.

[12] Y. C. Chen, D. Bakavos, A. Gholinia and P. B. Pragnell, "HAZ development and accelerated post-weld natural ageing in ultrasonic welding aluminium 6111-T4 automotive sheet," *ActaMaterialia*, 60, pp. 2816–2828, 2012.

[13] D. Bakavos and P. B. Pragnell, "Mechanism of joint and microstructure formation in high power ultrasonic spot welding 6111 aluminium automotive sheet," *Material Science and Engineering A*, 527, pp. 6320-6334, 2010.

[14] M. Shakil, N. H. Tariq, M. Ahmed, M. A. Choudhary, J. I. Akhter and S. S. Babu, "Effect of ultrasonic welding parameters on microstructure and mechanical properties of dissimilar joints," *Materials and Design*, 55, pp. 263-273, 2014.

[15] J. Magin and F. Balle, "Solid state joining of aluminum, titanium and their hybrids by ultrasonic torsion welding," *Mat.-wiss.u.Workstofftech*, 45(12), pp. 1072-1083, 2014.

[16] C. Jhang and L. Li, "A Coupled Thermal-Mechanical Analysis of Ultrasonic Bonding Mechanism," *Metallurgical and Materials Transactions B*, 40(B), pp. 196-207, 2009.

[17] C. Q. Zhang, J. D. Robson, O. Ciuca and P. B. Pregnell, "Microstructural characterization and mechanical properties of high power ultrasonic spot welded aluminum alloy AA6111-TiAl6V4 dissimilar joints," *Materials Characterization*, 97, pp. 83-91, 2014.

[18] C. Q. Zhang C. Q., J. D. Robson and P. B. Pregnell, "Dissimilar ultrasonic spot welding of aerospace aluminum alloy AA2139 to titanium alloy TiAl6V4," *Journal of Materials Processing Technology*, 231, pp. 382-388, 2016.

[19] P. Pragnell, F. Haddadi and Y. C. Chen, "Ultrasonic spot welding of aluminum to steel for automotive applications-microstructure and optimization," *Materials Science and Technology*, 27(3), pp. 617-624, 2011.

[20] Zhu, K. Y. Lee and X. Wang, "Ultrasonic welding of dissimilar metals, AA6061 and Ti6Al4V," *Int J Adv Manuf Technol*, 59, pp. 569-574, 2012.

[21] F. Balle and J. Magin, "Ultrasonic spot and torsion welding of aluminium to titanium alloy: Process, properties and interfacial microstructure," *Physics Procedia*, 70, pp. 846-849, 2015.

[22] S. Q. Wang, V. K. Patel, S. D. Bhole, G. D. Wen and D. L. Chen, "Microstructure and mechanical properties of ultrasonic spot welded Al/Ti alloy joints," *Materials and Design*, 78, pp. 33-41, 2015.

[23] F. Haddadi and F. Abu-Farha, "Microstructural and mechanical performance of aluminum to steel high power ultrasound spot welding," *Journal of Materials Processing Technology*, 225, pp. 262-274, 2015.

[24] Shin-ichi Matsuoka and H. Imai, "Direct welding of different metals used ultrasonic vibration," *Journal of Material Processing Technology*, 209, pp. 954-960, 2009.

[25] Hiromichi T. Fujii, Y. Goto, Y. S. Sato and H. Kokawa, "Microstructural evolution in dissimilar joint of Al alloy and Cu during ultrasonic welding," *Materials Science Forum*, 783, pp. 2747-2752, 2014.

[26] Hiromichi T. Fujii, Y. Goto, Y. S. Sato and H. Kokawa, "Microstructure and lap shear strength of the weld interface in ultrasonic welding of Al alloy to stainless steel," *Scripta Materialia*, 116, pp. 135-138, 2016.

[27] Y. Y. Zhao, D. Li, and Y. S. Zhang, "Effect of welding energy at interface zone of Al-Cu ultrasonic welded joint," *Science and Technology of Welding and Joining*, 18 (4), pp. 354-360, 2013.

[28] C. Y. Zhang, D. L. Chen and A. A. Luo, "Joining 5754 Automotive Aluminum Alloy 2-mm Thick Sheets Using Ultrasonic Spot Welding," *Welding Journal*, 93, pp. 131s-138s, 2014.

[29] Xin Wu, Tang Liu and Wayne Cai, "Microstructure, welding mechanism and failure of Al/Cu ultrasonic welds," *Journal of Manufacturing Processes*, Vol. 20, pp. 321-331, 2015.

[30] R. Jahn, R. Cooper and D. Wilkosz, "The Effect of Anvil Geometry and Welding Energy on Microstructures in Ultrasonic Spot Welds of AA6111-T4," *Metallurgical and Materials Transactions A.*, 38(A), pp. 570-583, 2007.

[31] R. Balasundaram, V. K. Patel, S. D. Bhole and D. L. Chen, "Effect of Zinc interlayer on ultrasonic spot welded aluminium to copper joints," *Material Science and Engineering A*, 607, pp. 277-286, 2014.

[32] Panteli, J. D. Robson and P. B. Pragnell, "The effect of high strain rate deformation on intermetallic reaction during ultrasonic welding aluminum and magnesium," *Materials Science & Engineering A*, 556, pp. 31-42, 2012.

[33] V. K. Patel, S. D. Bhole and D. L. Chen, "Influence of ultrasonic spot welding on microstructure in a magnesium alloy," *Scripta Materialia*, 65, pp. 911-914, 2011.

[34] V. K. Patel, S. D. Bhole and D. L. Chen, "Formation of zinc interlayer texture during dissimilar ultrasonic spot welding of magnesium and high strength low alloy steel," *Materials and Design*, 45, pp. 236-240, 2013.

[35] V. K. Patel, S. D. Bhole and D. L. Chen, "Improving weld strength of Magnesium to Aluminium dissimilar joints via Tin interlayer during ultrasonic spot welding," *Science and Technology of Welding and Joining*, 17(5), pp. 342-347, 2012.

[36] T. Watababe, H. Sakuyama and A. Yanagisawa, "Ultrasonic welding between mild steel sheet and Al-Mg alloy sheet," *Journal of Materials Processing Technology*, 209, pp. 5475-5480, 2009.

[37] A. Macwan and D. L. Chen, "Microstructure and mechanical properties of ultrasonic spot welded copper-to-magnesium alloy joints," *Materials and Design*, 84, pp. 261-269, 2015.

[38] Ibrahim E. Gunduz, T. Ando, E. Shattuck, P. Y. Wong and C. C. Doumanidis, "Enhanced diffusion and phase phase transformations during ultrasonic welding of zinc and aluminium," *Scripta Materialia*, 52, pp. 939-943, 2005.

[39] M. P. Sathpathy, B. R. Moharana, S. Dewangan and S. K. Sahu, " Modeling and optimization of ultrasonic metal welding on dissimilar sheets using fuzzy based genetic algorithm approach," *Engineering Science and Technology, an International Journal*, Vol. 18, pp. 634-647, 2015.

Wireless Networks
https://doi.org/10.1007/s11276-018-1750-z

New scheme for underwater acoustically wireless transmission using direct sequence code division multiple access in MIMO systems

Rajiv Kapoor[1] · Rashmi Gupta[2] · Raghavendra Kumar[3] · Le Hoang Son[4,5] · Sudan Jha[6]

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
This paper proposes a new technique based on Direct Sequence Code Division Multiple Access for underwater acoustically wireless transmission with excessive transmission rate. Environment of subsea is challenging for wireless communication because the medium in which waves are propagating is not air. In fact, it is propagated through fractions of water having different densities. Finding out various techniques for multipath access targeting the physical layer of Acoustic Sensor Networks is indeed necessary. The recent approaches have suggested that coded modulation techniques with exploited diversity are highly preferred in order to enhance the dependability of the acoustic link in different multipath channels. The proposed technique divides the channel into sub ones and transmits information via those sub channels. In variety-spectrum, a signal in a bandwidth is unfold within frequency domain and broad bandwidth. Experimental results show that Bit Error Rate (BER) of this method is better than that of channel equalization in the respective systems.

Keywords Magnetic · Induction · Wireless communication · Channel model · Underwater wireless communication networks (UWCNs) · Underwater acoustic communications · CDMA · Bit error rate (BER)

1 Introduction

In the past decades, the most widely used communication technique for the *Underwater Wireless Communication Network* (UWCN) was *acoustic communication* defined in the physical layer for long distance communication [1]. However, it has poor performance in shallow water and other disadvantages such as air bubbles in water and low temperature gradients due to speed of transmission [2]. Another technique namely *Optical underwater communication* has high bandwidth (gigabits per second) and low manufacture cost; yet it still remains disadvantages in line-of-sight communication and tight alignment of nodes [3]. This type of communication does not cross water/air boundary and also affects particles, marine fouling, soil and rock. *Electromagnetic wireless communication* is the third option with high bandwidth at very short range [4]. Nevertheless, it suffers from high path loss and fading as well as having large antenna size and dynamic channel condition [5].

The radio waves propagating requires high scaled antennae with huge transmission power, subjected to the condition that they are passed through the salted water at

✉ Le Hoang Son
sonlh@vnu.edu.vn

Rajiv Kapoor
Rajivkapoor.dtu@gmail.com

Rashmi Gupta
rashmig71@yahoo.com

Raghavendra Kumar
raghvendraagrawal7@gmail.com

Sudan Jha
jhasudan@hotmail.com

1 Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

2 Department of Electronics and Communication Engineering, AIACT&R, Delhi, India

3 Computer Science and Engineering Department, LNCT College, Bhopal, MP, India

4 Institute of Research and Development, Duy Tan University, Danang, Vietnam

5 VNU University of Science, Vietnam National University, Hanoi, Vietnam

6 School of Computer Engineering, KIIT University, Bhubaneswar, India

Published online: 12 May 2018

extremely degraded frequencies ranging from 30 to 300 Hz. Indeed, these salty water remains highly conductive [4]. Code Division Multiple Access (CDMA) suits well in direct sequence spread-spectrum modulation. If we consider the same for mobile underwater wireless communication networks, CDMA has proven to be the best solution so far [6]. The reason is a specific frequency selection for fading channels which are highly vigorous but subjected in compromising with the multipath effect at the receiver's end by using manipulated Rake filters. The noted advantage of this approach is that we can collect various sources of energies that are transmitted and dispersed through multiple rays thus allowing receivers to identify and recognize signals in various multiple devices [7]. However, Underwater Wireless Acoustic Communication (UWAC) possesses very high confining conditions than terrestrial radio communication. When the signal bandwidth is limited, the enhancement of communication capacity becomes tough because this technique utilizes a degraded frequency in comparison to the terrestrial radio communication. This is due to media characteristics [8].

Submerged acoustic channel is thought to be a good remote medium [9]. The unpredictability of submerged acoustic channels is led by ocean condition qualities which fuse colossal postponement, twofold side-spreading, doppler-spreads, recurrence particular blurring, and obliged data transfer capacity [10]. Notwithstanding, productive submerged interchanges are basic to many sorts of logical and common missions in the sea (e.g. sea checking, sea investigation, undersea save, and undersea debacle reaction). Together with sensor advancement and vehicular development, remote submerged exchanges are appealing to engage new applications from natural seeing to get-together of oceanographic data, marine antiquarianism, and chase and spare missions [11]. As electromagnetic waves engender ineffectively in ocean water, acoustics gives the most evident medium to empower submerged correspondences. Correspondence systems initially created for earthbound wired and wireless channels require huge changes to suit submerged channels [11].

New advances of rapid correspondence for picture and video transmission are likewise attractive to encourage the up and coming era of effective undersea campaigns [12]. Current acoustic correspondence advancements can give constrained information rates because of the specific physical highlights of channel [8]. The relating wireless innovation for undersea interchanges still needs huge further improvement [8]. Interestingly, acoustic waves can spread over longer separation. A submerged acoustic channel shows a correspondence framework creator with numerous challenges. The properties of the submerged medium are likewise to a great degree fluctuated, and change in both space and time [12]. Vacillation is from natural qualities incorporate occasional changes, geological varieties both in temperature and saltiness, seabed alleviation, and streams. This makes the submerged acoustic flag fluctuating arbitrarily. In that capacity, the choice of adjustment and mistake revision systems is exceptionally testing. To alleviate the bandwidth constraint, multicarrier adjustment is utilized as a part of the submerged acoustic channel. It is another option to conquer the long-lasting delays in submerged acoustic channels [6].

Underwater Wireless Acoustic Communication with Multiple-Input Multiple-Output (MIMO) is mostly used in underwater channels in which several antennas are utilized at the both ends of transmission (usually in receiver) and transmitter two or more antennas are outfitted [13, 14]. System's spatial diversity and multiplexing defines these systems by using *tx* and *rx* ends. Space Time coding technique is used in a MIMO link to enhance reliability [9]. Redundancy is integrated in the transmitter so that forward error correction at the receiver will retrieve exact sent data [15]. This technique contains two groups: *Space Time Block Codes (STBC)* aims to build an arrangement of orthogonal code-words which are sent along the antennas; *Space Time Trellis Codes (STTC)* disperses a trellis code among multiple antennas subsequently providing coding and diversity gain [16]. The receivers depend upon the Viterbi algorithm to decode the information, since the trellis coding method is convolution, thus expanding the system complexity [17]. Some other methods namely the Direct Sequence Code Division Multiple Access (DS-CDMA) [18, 19], Orthogonal Frequency Division Multiplexing (OFDM) [12, 19–21], MIMO [22], Modulation Errors Correction [23] and others have not assured the absolute performance in terms of communication beneath water [24, 25]. Hence, it is necessary to investigate a new approach and technique to find better performance [7].

In this paper, we propose another strategy for acoustic submersion and remote transmission with extraordinary transmission rate based on DS-CDMA and MIMO. The procedure is veritable layer and packages the channel into sub ones, and transmits data for those sub channels. The codes utilized are pseudo-sporadic tumult groupings. In a mix expand procedure, a flag conveyed in a picked data trade constrain is deliberately spread out inside the rehash space, understanding a standard with an essentially wider transmission restrain. The proposal is presumably supported.

The structure of the paper is followed by discussion of the related works in Sect. 2. Section 3 analyzes underwater channel models for communication. Section 4 proposes the new model. Section 5 shows experiments followed by conclusions in Sect. 6.

2 Related works

Herein, we analyze in details the CDMA approaches. The existing Mac outlines are all around whirled around Transporter-Sense-Organized-Accessor CDMA. The reason is the theory of Repeat Division Particular Access, which is declared as unauthentic for UW-ASNs in light of the compelled information transmission in Submerged Acoustic (UW-A) channels and shortcoming of obliged band structures to scattering and multipath [26, 27]. Time Division Changing Access (TDMA) shows obliged exchange speed adequacy by uprightness of long-time screens required in UW-A channels. The variable concede makes it to a mind-boggling degree hard to get a handle on a correct synchronization [1]. Discussions showed CDMA being the most encouraging physical layer and different access system [18].

In [28], two spread-run physical layer systems for comprehension, in particular Direct Strategy Spread Range (DSSS) and Repeat Ricocheting Spread Range (FHSS), are considered that shallow the water trades. While in DSSS, data are spread to keep the commonplace impedance, and on the other hand in FHSS, synchronous exchanges use particular repeat skipping groupings; therefore transmitting on different repeat social gatherings. Inquisitively in the submerged condition, FHSS prompts a higher piece misuse rate than DSSS [28]. Another attracting get to system mixes DSSS-CDMA with multi-conveyor transmissions [29], which may offer higher not well characterized vision efficiency than its single-transporter relate. Thus, data rate which are very high can be kept up by developing the length of every photograph, diminishing Cover Picture Avoidance. In any case, multi-transporter sending is not sensible for sensors which are low-end by virtue of high multifaceted nature. It is required to focus on single-transport CDMA to keep the eccentricity of favored viewpoint restricted sensor handsets lower. More importantly, the above studies in a general sense considered CDMA from the view of physical layer, i.e., the sensibility of different sorts of CDMA-based transmission structures.

Submerged acoustic high-rate correspondence plots an influencing change to perspective that attestations to enable or overhaul a couple of key applications in oceanic research, for instance, data gathering contamination checking, key observation and debacle question. In submerged structures, acoustic exchanges will be used as radio waves affect insufficiently in water. Outline of information transmission valuable coded alter plans, in setting of low-bizarreness, non-iterative unraveling, requires the data of the channel obscuring bits of learning. In any case, submerged acoustic (UWA) channels indicate time-changing obscuring estimations, which depend on a combination of factors, including the sea state, ocean vitality, and the growth way length. Rayleigh blurring models are proposed in [19]; of course, Rician obscuring models are in [18]. These models recognize an expansive number of scrambles considering the utilization of extremely far hypothesis. STBC diagrams has gotten in all cases figured which can finish full transmit better than anything typical mix and a higher code rate, however requiring basically encourage unraveling abnormality, are flabbergasting.

Remote correspondence structures exhibit a basic change the degree that degree and execution if remote exchanging frameworks are used [19]. Raise and forward (AF) is the most enrapturing approach among each and every fantastic custom by virtue of its fundamental organizing at the hand-off terminals. A huge interest is gotten by various input–diverse yield producing frameworks by virtue of their exceptional conditions, for instance, widened confine, data rate, gathered mix get, and demonstrate get. A key doubt in by a wide edge most by far of the wears out MIMO radiating is due to the prompt CSI which is staggeringly known at the exchange and target centers [12, 18, 20, 21]. Starting late, the hostile goof execution of a singular party mechanical social event based astute AF exchange system has been investigated in [22] by using a fell channel estimation tradition. To the best of our understanding, the true-blue execution of AF MIMO exchange structures, utilizing orthogonal space– time block code (OSTBC) transmission and assessed channel cross sections, isn't analyzed so far in the creation.

Poor and Tanda [30] managed the issue of multiuser ID in deal with strategy CDMA blurring paths with rash perplexity. Along these lines, M-estimator-based structure for non-sensible demodulation of differentially sort out move keyed (DPSK) signals transmitted meanwhile by systems for a CDMA level obscuring derail installed in surged pandemonium, is proposed and isolated. Adjoudani et al. [11] suggested a multiple-input–multiple-output (MIMO) augmentation for a third-era (3G) wireless framework. The combination of MIMO ideas within the current UMTS standard and the related space–time RAKE recipient are clarified. Liu et al. [31] surveyed the physical basics and designing usage for productive data trade by means of wireless correspondence utilizing physical waves as the bearer among nodes in a submerged sensor organize. The physical waves under talk incorporate sound, radio, and light. Jiang et al. [32] studied the current research improvement of Underwater Acoustic Networks (UANs). It investigates the uniqueness of submerged acoustic channel first. A few down to earth issues of UANs are then raised, running from organize topology, control effectiveness, physical layer, MAC layer, arrange layer to application layer. Manjula et al. [10] presented a few essential key angles and models of UWSNs, rising examination issues of

submerged sensor systems and uncovered the scientists into systems administration of submerged specialized gadgets for energizing sea checking and investigation applications. Tao et al. [33] suggested another time-area MIMO equalization strategy for single transporter UWA correspondences. Others studies and applications can be seen in [34–40, 41–43].

3 Underwater acoustic channel

3.1 Characteristics of underwater acoustic channel

UWAC happens to be, ever known, the most complex communication channel. It conveys numerous challenges to the underwater communication. Its characteristics consist of the frequency bandwidth limitation, varying multipath, high level of noise and fast fading. The diffusion and the absorption of acoustic energy give an account of frequency and transmission distance. Thus increase in frequency and transmission distance increases transmission loss. This feature causes immense attenuation for long distance transmission of high frequency signals. The authors [2] showed that in the near distance (i.e. < 1 km) bandwidth can be more than 100 kHz; in the middle distance (i.e. 1–20 km) the bandwidth is restricted in magnitude of 10 kHz; in long distance (i.e. 20–2000 km) the bandwidth is about a few kHz only. Ocean consists of many noise sources. Certain noise sources (tide and wave caused hydrostatic pressure's effect; disturbance due to earthquake; sailing ship; thermal noise; surface wave; on flow) are noted down with the rise in frequency [5]. For 10 kHz range of frequency, high noise makes the recovery of the original carrier complex.

Reflection at the surface and at the base of ocean, what's unbendingly considered to be the territory of diffuses and reflectors, influence acoustic wave to satisfy the recipient along 2 or 3 specific courses after transmission. This consideration is named as multipath transmission. It is an essential magnificent degree factor that effects the submerged acoustic correspondence execution. The multipath transmission causes hail turning (stunning jumbling) and particular obfuscating. The stage and the abundance of the standard will change with repeat and time which makes getting part fouls up. To manage these issues, change approach, set away thinking, spread range framework and better than anything standard social gathering structure need to be thoroughly looked upon.

3.2 Model of submergedacoustic channel

Submerged channels can be adjusted by arranging time-varying channel [28]. Considering multipath transmission in the condition, underneath Fig. 1 tends to the model.

Considering the input signal as $q(t)$, representation for output signal is:

$$R(t) = \sum_{k=1}^{L} a_k(t) q(t - \tau_k(t)) + w(t) \quad (1)$$

where L represents number of multipath; $a_k(t)$ represents time-varying attenuation factor in k_{th} transmission path; $\tau_k(t)$ represents delay; $w(t)$ represents additive white noise.

3.3 Review of the model/protocol

Right when focus centers are passed around, courses will be made in the instatement deal with. Each inside point contained by a way is asked for as appeared by its position in that way, which recommends, the standard concentration point in a path changes into the fundamental sender, however the running with center point in that course changes into the second sender and whatnot. The last objective is the sink. The ask for is trailed by each inside point in an approach to manage perform transmission as showed by the round-robin way. At a minute that a center guide needs toward transmit a package, stand-out event of the basic sender, it needs to sit tight for clusters from its past concentration point.

In Fig. 2, at the transmission time the process goes on as per the following: node A, the earliest sender, transmits the corresponding packet (packet 1) to the next nearest hop node, here node B. After receiving the node A's packet 1, node B forwards this and then sends the packet corresponded to itself (packet 2). Node B sends packet 1 back to the node A, until that node A does not delete packet 1. After receiving packets from node B i.e. packet 1 and 2, node C forwards them to node D and transmits packet 3 which is a short passing token only, because it does not have any data to transmit. This short passing token triggers the next transmission. After receiving packets from node C i.e. packet 1, 2 and 3, node D forwards them and transmits packet 4 to the sink.

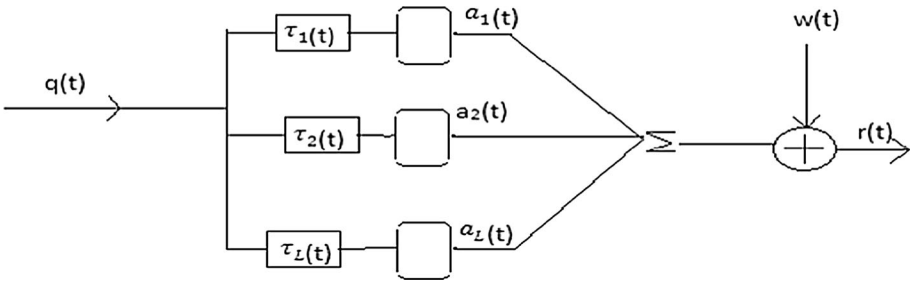


Fig. 1 Time-changing multipath with AWGN

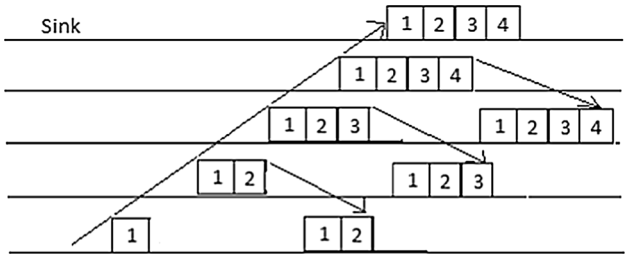


Fig. 2 Transmissions of packets by means of a well set up path

4 Proposed methodology

4.1 Algorithm at the receiver's end

4.1.1 System model

The expression for transmitted signal, which is complex baseband, for *i*th user can be written as:

$$z_i(t) = \sum_n c_i(n) j_i(t - nT), \quad i = 1, \ldots .I \tag{2}$$

Here $c_i(n)$ is representing the contiguous sets of data, $j_i(t)$ is representing the distributed waveform and T gives the symbol period. Values of the data symbol may be obtained from a discretionary Phase Shift Keying (i.e. PSK) or Quadrature Amplitude Modulation (QAM) alphabet. The expression for spreading waveform can be written as:

$$j_i(t) = \sum_{l=0}^{L-1} p_i(l) j(t - lT_c) \tag{3}$$

Here $p_i(l)$ represents sequence of chip for *i*th code, $j(t)$ is the transmitter shaping pulse and T_c is the interval between the chips. The time duration of code is equivalent to the duration of the symbol i.e. $LT_c = T$. The chip sequence $p_i(l)$ either may be real-valued ($\pm$ 1) or complex. Alternatively, the transmitted signal is,

$$z_i(t) = \sum_k q_i(k) j(t - kT_c) \tag{4}$$

$$q_i(nL + l) = c_i(n) p_i(l), \quad l = 0, \ldots L - 1, \forall n \tag{5}$$

Frequencies from various clients processing speed is fast to unmistakable and multipath-distorted methodology and reach at a typical recipient. Two receivers are being depicted underneath; one depends upon picture choice info while substitute depends upon chip speculation criticism.

4.1.2 Receiver's end

Let us drop its index I for notation simplicity, that is $p(l) = p_i(l)$ and $c(n) = c_i(n)$. The received complex base-band signal in k^{th} element is represented by $v_k(t)$. Sampling is performed on these signals at N_y having N_s samples for T_c. All K received signals are combined to form the inputs $P \leq K$ for feed forward filters organized into column vectors c_p of K-elements. Forward filters are placed at T_c/N_s fractionally where each has N coefficients in a_p. kth receiver's temporal samples are organized $\text{inv}_k(n, l) = v_k(nL + l)$. (n, l) represents the lth chip of the nth data symbol. Vectors $v_k(n, l)$ consist of nth transmitted data symbol $d(n)$. The following matrix is formed by signal vectors from all receive elements:

$$V(n, l) = V(nL + l) = \begin{bmatrix} v_1^T(n, l) \\ \vdots \\ v_K^T(n, l) \end{bmatrix}, \quad l = 0, \ldots L-1 \tag{6}$$

Output signals of feed forward equalizers are synchronized by $\hat{\theta}(n)$. In most applications, a single carrier phase is sufficient to achieve good computational efficiency. The feedback filter consists of M coefficients in $\boldsymbol{b}$. The nth data estimation is obtained as,

$$\hat{c}(n) = \frac{1}{L} \sum_{l=0}^{L-1} c_p^{'} V(n, l) a_p^* e^{-j\hat{\theta}(n)} - b' \tilde{c}(n - 1) \tag{7}$$

Prime is denoting the conjugate transpose here. In this expression, dependence of the data is shown upon receiver parameters, i.e. phase estimate, feed forward filters, combiner weights and feedback filter. For minimizing the MSE in the data detection all these parameters are altered jointly. A single error drives the adaptive algorithm,

$$e(n) = \tilde{c}(n) - \hat{c}(n) \tag{8}$$

4.1.3 CDMA detectors for both Single and multi-users

UW-Mac (Submerged Mac) is transmitter-based Direct Amassing CDMA (DS-CDMA) method of reasoning for UW-ASNs which execute a novel close float passed on acknowledging to generally sort the code length and the perfect transmit centrality to compensate the neighboring far effect. A multi-customer identifier may be used by UW-Mac on asset rich gadgets, for example, surface stations, AUVs and UW-segments and a specific customer pioneer on low-end sensor centers. The information bearing signed, in DS-CDMA correspondence systems, is energetic imitated by a spreading code which has a more discernible transmission confine than the data itself. The recipient covers the spread range signal, which was transmitted, utilizing an unobtrusively made code gathering. The recipient, for the covering operation reason, must know the code gathering used to spread the standard. Additionally, the obscurely made code and the got hail must be really synchronized. Toward the starting time of the party this synchronization must be refined and spared until the minute that the moment that the entire pennant has been gotten completely. In a DS-CDMA chart the principal issue experienced is the Multiuser Access Impedance (MAI), which is helped by synchronous transmissions from unmistakable customers. Genuinely, as far as possible is obliged by the establishment perplexity just [44]. In this way, to reduce the MAI, low cross-relationship in between the pined for customers and the infringing customers is crucial. In like way, for time endeavored basic synchronization, satisfying auto-association properties are required. Shockingly, both autocorrelation and cross-relationship properties can't be in the meantime redesigned.

4.2 The MIMO channel

A mathematical approach is required to comprehend MIMO's principles. As in Fig. 3, every transmitter antenna adds to the received signals in every receiver. We will consider a flat fading, narrow band channel for a channel model and the system expressions will be written in the frequency domain. The received signal for an antenna, for instance, is being expressed as:

$$y_{f_o}^r = \sum_{t=1}^{M_t} H_{f_o}^{t_r} d^t + w_k^r r = [1 \ldots M_r] \tag{9}$$

Here **t** and **r** depicts the transmitter and receiver index respectively and $\mathbf{f_0}$ depicts the frequency of operation. For notation simplicity, time index has been dropped. $\mathbf{M_t}$ and $\mathbf{M_r}$ represent the number of transmitters and receivers used in the system respectively.

We rewrite the previously written equation in the matrix notation for clarity. Define the vectors:

$$y_{f_o} = \left[y_{f_0}^1, \ldots\ldots, y_{f_o}^{M_r}\right] \tag{10}$$

$$c = [c^1, \ldots., c^{M_r}] \tag{11}$$

$$w_{f_o} = \left[w_{f_0}^1, \ldots\ldots, w_{f_o}^{M_r}\right] \tag{12}$$

$$\sim H_{f_0} = \begin{bmatrix} H_{f_0}^{1,1} & \cdots & H_{f_0}^{M_t,1} \\ \vdots & \ddots & \vdots \\ H_{f_0}^{1,M_r} & \cdots & H_{f_0}^{M_t,M_r} \end{bmatrix} \tag{13}$$

Now, the expression given above in Eq. (10) becomes:

$$y_{f_o} = \tilde{H}_{f_0} d + w_{f_0} \tag{14}$$

While expressing the development obtained in unessential packaging, examination of the condition should be possible. From condition (14), we understand that every transmitter adds to each and every one of the recipients and that certain terms and conditions on the channel structure $\tilde{H}_{f_0}$ will be major to demodulate the basic standard d. Interface with us to take a nearer see at the physical, true blue

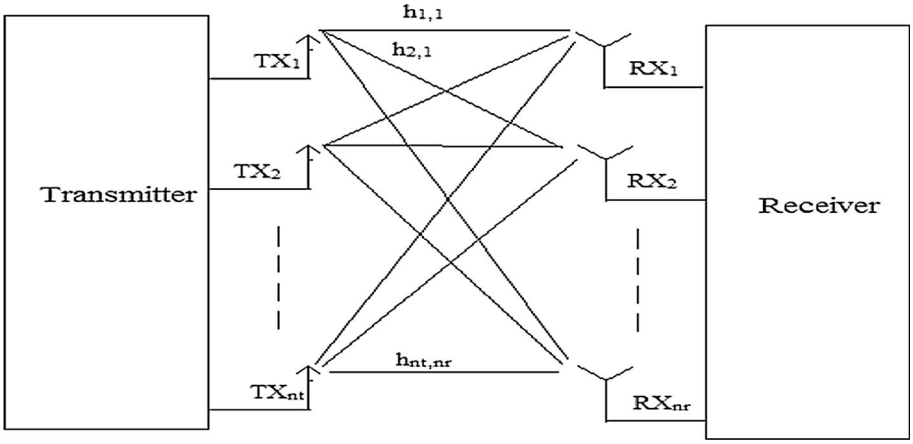


Fig. 3 Simplified scheme of the MIMO channel

delineation of the condition appeared above, rather than inquiring about the demodulation system. Concerning SISO structures, Shannon has totally done; a nearby way channel explanation for containment can be supervised for MIMO frameworks. The key principles are the same: the inspiration driving detainment will be the most brilliant customary data between the information starting at now the channel, *d* and the information got, *y*. A couple of examinations are as following:

1. If there is a deterministic channel, the value which was expected is not utilized. If there is a random channel, with the use of average values, ergodic capacity is utilized.
2. R_{dk} matrix correlates between the data transmitted and fundamentally, on the transmitters, this matrix provides the power allocation.
3. Only at the transmitter side, channel capacity becomes useful. Hence, CSI or channel state information is required at the transmitter side.
4. If there is no CSI available then no strategies can be performed on power allocations. Then, in each of the antennas same measure of power will be transmitted.
5. To maximize the capacity, the methodology to be used depends upon the kind of fading expected for channel. Nakagami, Rician and Rayleigh channels models [45] are utilized in typical situations.

For a MIMO channel, huge number of studies by the researchers and specialists are been worked out and is still ongoing. This document will mainly focus on the demodulation strategies that could be applied rather than the statistical characterizations of an underwater acoustic MIMO channel. Now, since in STTCs is complexity much high, in a more detailed manner only STBC will be explained. Alamouti Coding [46] was utilized in the system outfitted in SPACE'08 experiment. Where, in each of the time block, we receive M_t symbols which can be formed into a matrix. This matrix does also create a space time symbol. Now this space time symbol helps us to define the coding type:

$$C = \begin{bmatrix} c_0^1 & \cdots & c_{N-1}^1 \\ \vdots & \ddots & \vdots \\ c_0^{M_t} & \cdots & c_{N-1}^{M_t} \end{bmatrix} \tag{15}$$

where c_{ij} gives the symbol sent at the time instant *j*, on the transmitter *i*. To make the set of code-words i.e. d_{ij}, 8i orthogonal among each other, is the primary purpose of the coding. The linear, simple and optimal decoding is outcome of this at the receiver. Like many other coding techniques, the system has to sacrifice the data rate because of redundancy. The number of encoded pictures reflects the rate of the code and is represented in a solitary time piece (observe that various transmitters can send a comparable picture) isolated by the amount of timetable openings essential to complete the space-square symbol, r = (number of symbols)/N. With a matrix of 2X2, the Alamouti's code is the simplest among such codes. For this code matrix is given below:

$$C = \begin{bmatrix} c_1 & c_2 \\ -c_2^* & c_1^* \end{bmatrix} \tag{16}$$

We clearly see that matrix columns are essentially orthogonal. Using optimal decoding technique, the BER of the Alamouti's code, at the receiver side, is equal to a MRC consisting 2 symbols over M_r number of receivers. This is a postponed outcome of ideal orthogonally between the photos after get dealing with: there are two duplicates of each photo. Considering that the photos can't be recouped after 2 openings of time (N for a standard STBC) along these lines presenting a little time delay. Despite the way that the commented setup was for STBCs, the STTCs are sturdier against messes up, however the beneficiary whimsy is higher in light of the way that dynamic programming estimations are required on the recipients for cure information recognizing evidence. In both ways, the utilization of STC engages MIMO structures to work with higher number of transmitters than recipients (Fig. 4).

5 Simulation results

We have gone through the application of CDMA-MIMI technology for underwater wireless communication. It is assumed that the channel outcomes are designed as Rayleigh fading. The realization of the channels is independent statistically from one to another user and from one to another block. We opted the Walsh-Hadamard code [47] which has excellent properties as the distributed code here (Table 1).

QPSK constellation [4] is used here for users data's mapping and de-mapping, and the least-square (LS) rule is embraced for the acoustic channel evaluation.

A MIMO-CDMA system consists of 2 transmitter antennas and 2 receiver antennas. The following are considered assumed for the simulation model:

1. The channel utilized is a gradually Rayleigh blurring, recurrence non-particular (i.e. level) channel and the between image obstruction (ISI), initiated by the channel, is absent there.
2. The channel gains of separate transmitter and receiver links are independent.
3. The perfect CSI information availability is assumed at the receiver side.

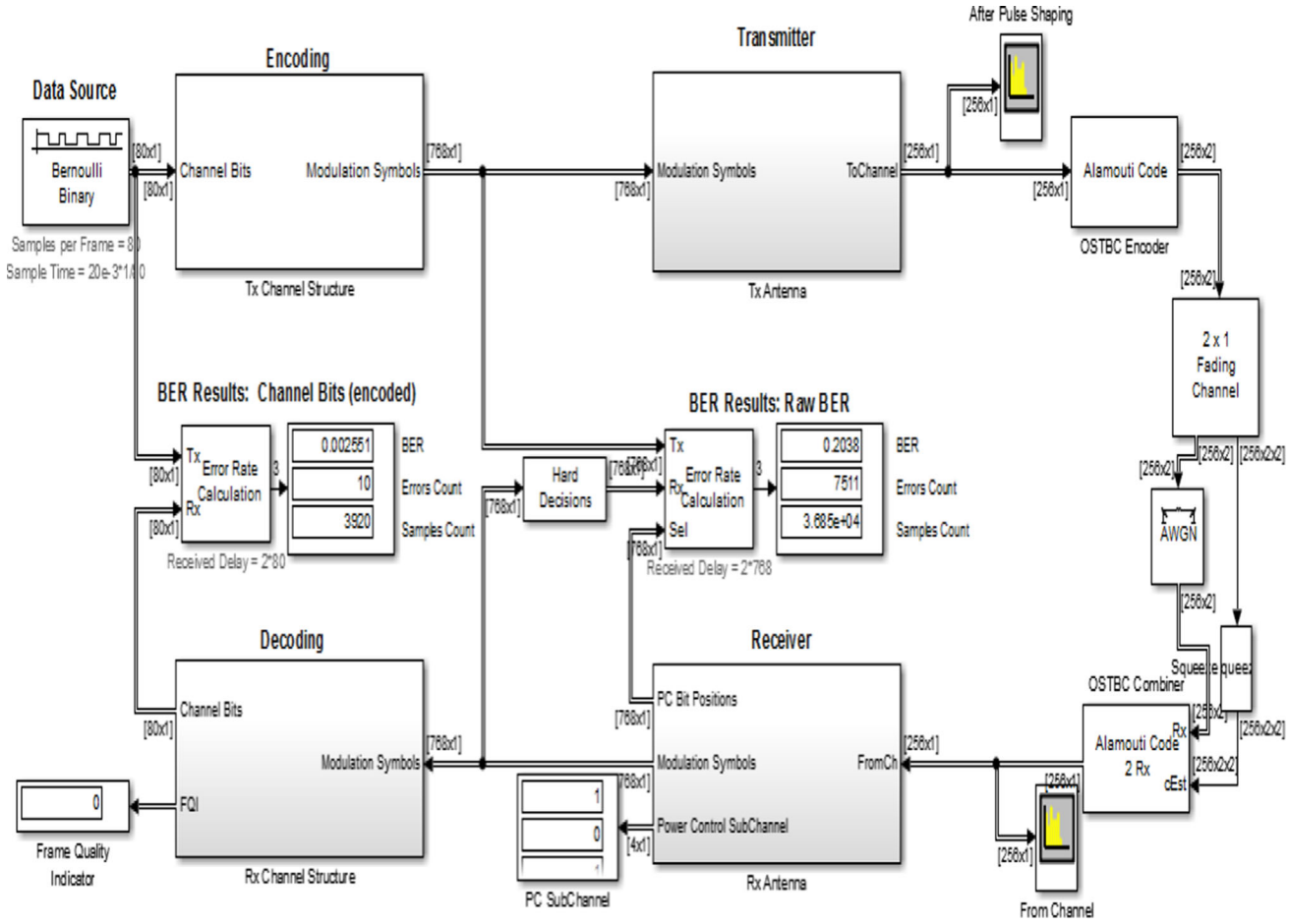


Fig. 4 Simulink model of CDMA-MIMO technique

Table 1 System specification for simulation

Spread code	Walsh–Hadamard
Walsh length	64
Over sampling rate	4
NFFT Size	257
Max Doppler frequency shift (pool)	10 Hz
Number of user	2
Max multipath delay (poor)	20 ms
Chip rate	1.2288e6
Channel spreading code index	23
Gate transmission rate	¼
PN-sequence offset	1

Figure 4 gives the MIMO-CDMA system's Simulink model. A 2 × 1 Rayleigh fading channel is used for a 2 × 2 MIMO system. It consists of 2 transmitter antennas and 2 receiver antennas. The estimated BER (Bit Error Rate) is 0.002551. The system parameters are shown in Table 1. On the basis of simulation, the performance of MIMO-CDMA is tested and analyzed in the UWA network. The lowest bit error rate (BER) is achieved here. Figure 5 represents the constellation diagrams at different stages. For each of these constellation diagrams we calculate the probability of error (P_e), signal energy efficiencies and performance of BER. The inputs are darkest ones in Fig. 5. Since the level of channel that is obstruction is missing in the next two images due to derotation, the outputs are slightly deemed and noises are not there with the output.

where MF stands for Matched Filter, DD stands for Decorrelating Detector, and M stands for estimation. Flag is reinforced through shaped channel or game-plan relating locator. Execution of the contraption under non-Gaussian test is unpalatable when showed up especially in relationship with that of Gaussian insurrection. M-estimation approach is utilized so that twofold getting wire structures are considered. Table 2 exhibits BER for all SNR respects in the scope of 15 and 30. Table 3 demonstrates the execution Relationship for one of kind estimators. Speculative charge is as bound and the sensible respect. Standard is

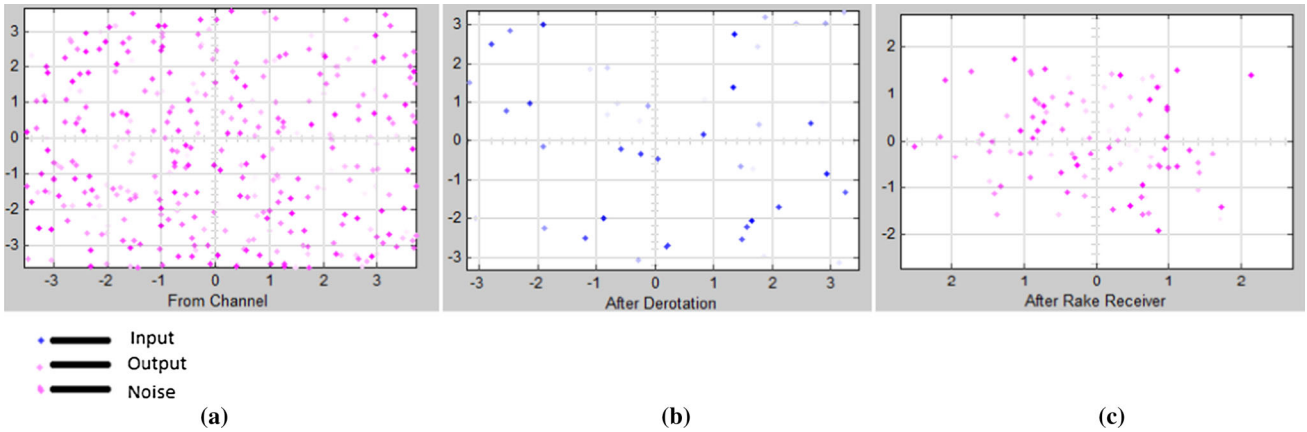


Fig. 5 Constellation diagram of **a** from channel **b** after derotation **c** after rake receiver

Table 2 Execution examination for various estimators with orthogonal spreading

		MF		DD		M-estimation	
Signal to noise ratio	Bit error rate (BER)	Gaussian noise (GN)	Non-GN	GN	Non-GN	GN	Non-GN
22	0.0016	0.0024	0.0053	0.0068	0.0123	0.0067	0.0074
23	0.0012	0.0033	0.0038	0.0053	0.0103	0.0052	0.0051
24	0.0010	0.0030	0.0063	0.0042	0.0011	0.0034	0.0029
25	0.0008	0.0019	0.0045	0.0051	0.0073	0.0027	0.0061
26	0.0006	0.0018	0.0033	0.0054	0.0113	0.0040	0.0064
27	0.0005	0.0010	0.0035	0.0061	0.0099	0.0025	0.0044

Table 3 Execution correlation for various estimators M-arrangement spreading

		MF		DD		M-estimation	
Signal to noise ratio	Bit error rate (BER)	Gaussian noise (GN)	Non-GN	GN	Non-GN	GN	Non-GN
22	0.0015	0.0034	0.0063	0.0048	0.0103	0.0087	0.0084
23	0.0014	0.0023	0.0048	0.0043	0.0113	0.0082	0.0061
24	0.0009	0.0020	0.0033	0.0052	0.0011	0.0074	0.0019
25	0.0007	0.0029	0.0045	0.0041	0.0063	0.0087	0.0051
26	0.0006	0.0028	0.0053	0.0044	0.0143	0.0080	0.0064
27	0.0045	0.0040	0.0055	0.0051	0.0079	0.0085	0.0054

sullied by speculation for jumbling and bewilder then as it comes the channel. Nonlinear expressive framework relating locators complete higher inside watching weaving change. In Table 3, we have built up a discrete-occasion supported package level test structure to review the execution of the proposed cross-layer convention. MIMO interface module produces bit mess up rate turns like information parameters (e.g. measures of transmit/get parts, decision of room time codes, demonstrate transmit control, acoustic tumult level, Doppler spread and relationship among various channels). Our submerged CDMA-MIMO module addresses a submerged acoustic relationship for various estimation of SNR.

We considered a CDMA-MIMO condition where settled length spreading code length 19 and two transmit and radio wires. It is expressed that central CDMA-MIMO improvements can be refined in acoustic channels. A coding structure for submerged acoustic channel gives a magnificent degree of execution.

We examine execution aftereffects of CDMA-MIMO and ALOHA in shallow water. It is noticed that all the above reenacted charts are accomplished by making

midpoints upon various topologies. The chip rate is 110 kcps, the dispersed code length is 18, the greatest transmission control Pmax is 12 W, the information parcel size is 240 Bytes, ITS, MTS, and ACK sizes are 12 Bytes. All sensors are arbitrarily conveyed in 3D shallow water of $500 \times 500 \times 50$ m^3. Figures 6, 7 and 8 assess CDMA-MIMO adaptability and versatility to channel crashes by fluctuating the quantity of sensors.

In Fig. 6, ALOHA is appeared to experience the abominable effects of more crashes and bundle retransmissions. In CDMA-MIMO, on a very basic level ITS and MTS can affect, and their size is humbler than the bundle study. This diminishment the impact likelihood, impelling a higher package transport rate. Figure 7 demonstrates the standard deferral of enough got packs. ALOHA can decrease the postponement of enough got packs without two-way handshake. Regardless, ALOHA drops an extensive measure of packs and its throughput is lower than that of CDMA-MIMO. In Fig. 8, CDMA-MIMO is appeared to unfathomably decrease the importance use (to not as much as half) by picking sensible transmit control. Likewise, ALOHA gobbles up an extensive measure of enormity in retransmitting social affairs. In Figs. 9, 10 and 11, we move the package between partition times to gage the

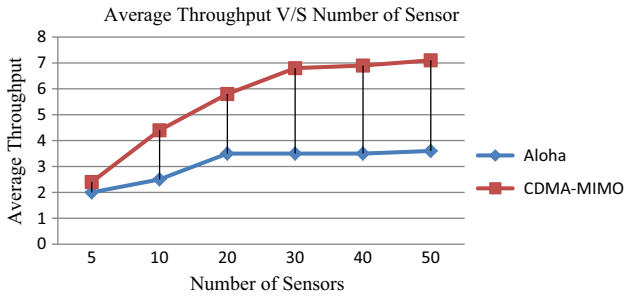


Fig. 6 Average throughput versus sensors' numbers

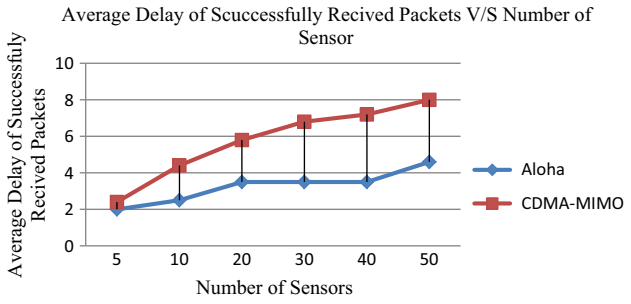


Fig. 7 Avg delay of successfully received packets versus number of sensors

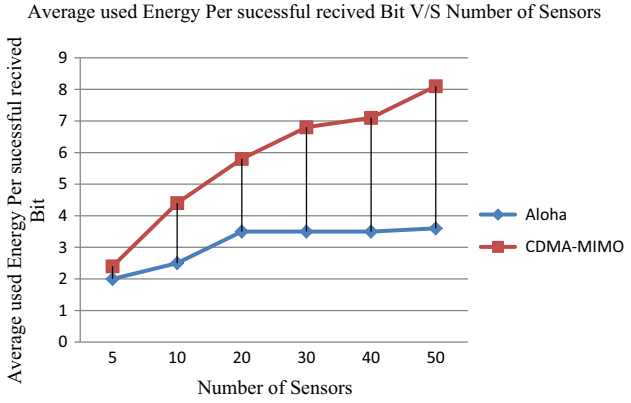


Fig. 8 Avg used energy per successful received bit versus number of sensors

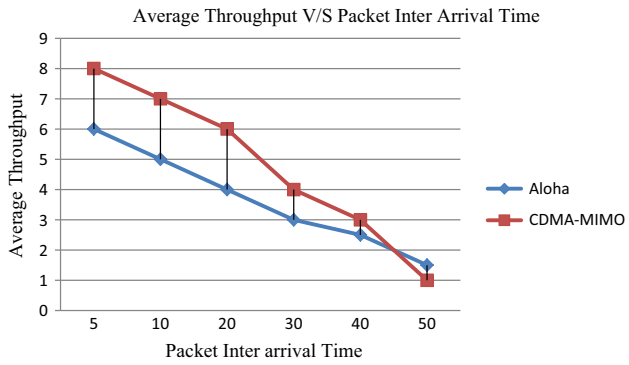


Fig. 9 Average throughput versus packet inter arrival time

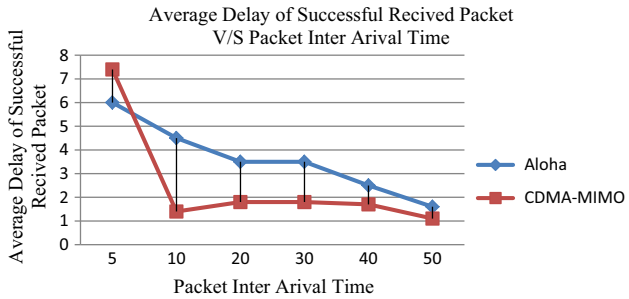


Fig. 10 Average delay of successful received packet versus packet inter arrival time

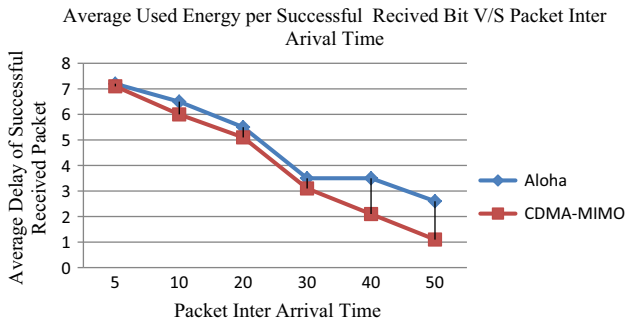


Fig. 11 Average delay of successful received packet versus Packet Inter arrival time

impact of movement. We set the bundle between area time to 20 s to keep up a key partition from line fortify surges. In Fig. 9, we set sensors' measure to 12 to evade a huge number of impacts due to different sensors getting to the path in the meantime. In Fig. 9, the CDMA-MIMO's throughput is higher than ALOHA even under overwhelming activity. Figure 10 shows that the delay of viably got packs is lower than that of ALOHA under overpowering movement conditions. Likewise, this grants CDMA-MIMO to save a great deal of essentialness as demonstrated in Fig. 11.

6 Conclusions

This paper proposed a new technique for underwater acoustically wireless transmission with excessive transmission rate based on Direct Sequence Code Division Multiple Access (DS-CDMA) and MIMO systems. In light for different submerged robots, this article considered a kind of submerged acoustic correspondence system with the likelihood of multiusers. With a specific end goal to guarantee the communication trustworthiness and quality, few key systems, e.g. channel coding and spread spectrum were embraced to decrease BER. The consequences of recreation show that this framework could effectively engage at most two customers. BER is almost 10–4 with the SNR of 15 dB, which is for all intents and purposes adequate. The results are significance in examination and investigation of submerged acoustic correspondence. However, this article does not consider the existing of carrier. Furthermore, the choice of PN code and the close far impact are additional issues. These are worth to be studied and researched in the future.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Pompili, D., Melodia, T., & Akyildiz, I. F. (2009). A CDMA-based medium access control for underwater acoustic sensor networks. *IEEE Transactions on Wireless Communications, 8*(4), 1899–1909.
2. Mulholland, J., & McStay, D. (2011). Wireless communication enhances subsea production monitoring. *Offshore, 71*(1), 76–77.
3. Lanbo, L., Shengli, Z., & Jun-Hong, C. (2008). Prospects and problems of wireless communication for underwater sensor networks. *Wireless Communications and Mobile Computing, 8*(8), 977–994.
4. Zhang, X., Cui, J. H., Das, S., Gerla, M., & Chitre, M. (2016). Underwater Wireless Communications and Networks: Theory and Application: Part 2 [Guest Editorial]. *IEEE Communications Magazine, 54*(2), 30–31.
5. Sun, Z., & Akyildiz, I. F. (2009). Underground wireless communication using magnetic induction. In *IEEE international conference on communications, 2009. ICC'09* (pp. 1–5).
6. Liu, L., Wang, R., Guo, D., & Fan, X. (2016). Message dissemination for throughput optimization in storage-limited opportunistic underwater sensor networks. In *2016 13th annual IEEE international conference on sensing, communication, and networking (SECON)* (pp. 1–9). IEEE.
7. Sozer, E. M., Proakis, J. G., Stojanovic, R., Rice, J. A., Benson, A., & Hatch, M. (1999). Direct sequence spread spectrum based modem for under water acoustic communication and channel measurements. In *OCEANS'99 MTS/IEEE. Riding the crest into the 21st century* (Vol. 1, pp. 228–233). IEEE.
8. Ziaeefard, S., Page, B. R., Pinar, A. J., & Mahmoudian, N. (2017). Effective turning motion control of internally actuated autonomous underwater vehicles. *Journal of Intelligent and Robotic Systems, 89,* 1–15.
9. Brady, D., & Preisig, J. C. (1998). Underwater acoustic communications. In H. V. Poor & G. W. Wornell (Eds.), *Wireless communications: Signal Processing perspectives* (Vol. 8, pp. 330–379). Upper Saddle River: Prentice-Hall Inc.
10. Singer, C., Nelson, J. K., & Kozat, S. S. (2009). Signal processing for underwater acoustic communications. *IEEE Communications Magazine, 47*(1), 90–96.
11. Chitre, M., Shahabudeen, S., Freitag, L., & Stojanovic, M. (2008). Recent advances in underwater acoustic communications and networking. In *OCEANS 2008* (Vol. 2008, pp. 1–10).
12. Manjula, R. B., & Manvi, S. S. (2011). Issues in underwater acoustic sensor networks. *International Journal of Computer and Electrical Engineering, 3*(1), 101.
13. Kojima, T., & Nagashima, Y. (2016). An initial acquisition scheme for Walsh-Hadamard code division multiplexing. In *2016 international conference on advanced technologies for communications (ATC)* (pp. 397–400).
14. Tam, N. T., Hai, D. T., Son, L.H., & Le Vinh, T. (2018). Improving lifetime and network connections of 3D wireless sensor networks based on fuzzy clustering and particle swarm optimization. *Wireless Networks*. https://doi.org/10.1007/s11276-016-1412-y.
15. Hai, D. T., Son, L. H., & Le Vinh, T. (2017). Novel fuzzy clustering scheme for 3D wireless sensor networks. *Applied Soft Computing, 54,* 141–149.
16. Jameel, F., Haider, M. A. A., & Butt, A. A. (2017). Performance analysis of VANETs under Rayleigh, Rician, Nakagami-m and Weibull fading. In *International conference on communication, computing and digital systems (C-CODE)* (pp. 127–132).
17. Son, L. H., & Thong, P. H. (2017). Soft computing methods for WiMax Network Planning on 3D Geographical Information Systems. *Journal of Computer and System Sciences, 83*(1), 159–179.
18. Chen, Y., Ji, F., Guan, Q., Wang, Y., Chen, F., & Yu, H. (2017). Adaptive RTO for handshaking-based MAC protocols in underwater acoustic networks. *Future Generation Computer Systems*. https://doi.org/10.1016/j.future.2017.08.022.
19. Hwang, S.-J., & Schniter, P. (2007). Efficient communication over highly spread underwater acoustic channels. *International Conference on Mobile Computing and Net-working, Montreal, 9–14,* 11–18.
20. Yang, R., Clement, B., Mansour, A., Li, M., & Wu, N. (2015). Modeling of a complex-shaped underwater vehicle for robust control scheme. *Journal of Intelligent and Robotic Systems, 80*(3–4), 491–506.

21. Iglesias, A. Song, J. Garcia-Frias, M. Badiey & G. R. Arce (2011). Image transmission over the underwater acoustic channel via compressive sensing. In *45th annual conference on information sciences and systems, Baltimore*, 23–25 March 2011 (pp. 1–6).
22. Tao, J., Zheng, Y. R., Xiao, C., Yang, T. C., & Yang, W. B. (2010). Channel equalization for single carrier MIMO underwater acoustic communications. *EURASIP Journal on Advances in Signal Processing, 2010*(1), 281769.
23. Jindal, H., Saxena, S., & Kasana, S. S. (2017). A sustainable multi-parametric sensors network topology for river water quality monitoring. *Wireless Networks*, 1–25.
24. Cui, J. H., Kong, J., Gerla, M., & Zhou, S. (2006). The challenges of building mobile underwater wireless networks for aquatic applications. *IEEE Network, 20*(3), 12–18.
25. Liu, S., Zhang, Z., & Yu, W. (2017). A Space-time coding scheme with time and frequency comb-like chirp waveforms for MIMO-SAR. *IEEE Journal of Selected Topics in Signal Processing, 11*(2), 391–403.
26. Singer, A. C., Nelson, J. K., & Kozat, S. S. (2009). Signal processing for underwater acoustic communications. *IEEE Communications Magazine, 47*(1), 90–96.
27. Costa, D., Palmieri, G., Palpacelli, M. C., Panebianco, L., & Scaradozzi, D. (2017). Design of a bio-inspired autonomous underwater robot. *Journal of Intelligent & Robotic Systems*. https://doi.org/10.1007/s10846-017-0678-3
28. Balanis, C. A. (2016). *Antenna theory: Analysis and design*. New York: Wiley.
29. Che, X., Wells, I., Dickers, G., Kear, P., & Gong, X. (2010). Re-evaluation of RF electromagnetic communication in underwater sensor networks. *IEEE Communications Magazine, 48*(12), 143–151.
30. Lanbo, L., Shengli, Z., & Jun-Hong, C. (2008). Prospects and problems of wireless communication for underwater sensor networks. *Wireless Communications and Mobile Computing, 8*(8), 977–994.
31. Su, X., Liu, X., Lu, X., & Luo, C. (2016). A combination of Alamouti code and beamforming technologies via dual-polarized antenna array systems. *International Workshop on Antenna Technology (iWAT), 2016* (pp. 119–122). IEEE.
32. Wang, Y., & Ding, F. (2016). Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. *Automatica, 71,* 308–313.
33. Kapoor, R., Gupta, R., Son, L. H., Jha, S., & Kumar, R. (2018). Detection of power quality event using histogram of oriented gradients and support vector machine. *Measurement, 120,* 52–75.
34. Balanis, C. A. (2012). *Advanced engineering electromagnetic*. New York: Wiley.
35. Francois, R. E., & Garrison, G. R. (1982). Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption. *The Journal of the Acoustical Society of America, 72*(6), 1879–1890.
36. Mishra, A. Anand, S., Ashmita, G., Padma, N., Padhi, P.P. & Kumar, J.M. (2010). Non-linear distortion analysis and computation of BER in BPSK OFDM systems. In *2nd national conference on recent trends in communication, computation and signal processing.*
37. Warsi, S., Jindal, V., Kumar, S., Koli, D., Bagchi, A., & Ribeiro, V. J. (2016). Joint scheduling and routing using space–time graphs for TDM wireless mesh networks. *Wireless Networks, 22*(7), 2355–2367.
38. Adjoudani, A., Beck, E. C., Burg, A. P., Djuknic, G. M., Gvoth, T. G., Haessig, D., et al. (2003). Prototype experience for MIMO BLAST over third-generation wireless system. *IEEE Journal on Selected Areas in Communications, 21*(3), 440–451.
39. Poor, H. V., & Tanda, M. (2002). Multiuser detection in flat fading non-Gaussian channels. *IEEE Transactions on Communications, 50*(11), 1769–1777.
40. Singh, K., Singh, K., Son, L. H., & Aziz, A. (2018). Congestion control in wireless sensor networks by hybrid multi-objective optimization algorithm. *Computer Networks, 138,* 90–107.
41. Hojo, K., Terada, Y., Yazawa, N., Watanabe, T., & Baba, T. (2016). Compact QPSK and PAM modulators with Si photonic crystal slow-light phase shifters. *IEEE Photonics Technology Letters, 28*(13), 1438–1441.
42. Jiang, Z. (2008). Underwater acoustic networks–issues and solutions. *International Journal of Intelligent Control and Systems, 13*(3), 152–161.
43. Hai, D. T., Tam, N. T., Son, L. H., & Vinh, L. T. (2016). A novel energy-balanced unequal fuzzy clustering algorithm for 3D wireless sensor networks. In *Proceedings of the seventh symposium on information and communication technology* (pp. 180–186). ACM.
44. Hai, D. T., Tam, N. T., Son, L. H., & Vinh, L. T. (2016). Experimental analyses of Terrain factors to performance of sensor placement optimization in 3D environments. *International Journal of Engineering and Technology, 8*(6), 394–400.
45. Liu, Q., Chen, X., Liu, X., & Linge, N. (2016). CACA-UAN: A context-aware communication approach based on the underwater acoustic sensor network. In *International conference on cloud computing and security* (pp. 37–47). Berlin: Springer.
46. Haci, H. (2017). Performance study of non-orthogonal multiple access (NOMA) with triangular successive interference cancellation. *Wireless Networks*. https://doi.org/10.1007/s11276-017-1464-7.
47. Li, X., Liu, J., Yao, Q., & Ma, J. (2017). Efficient and consistent key extraction based on received signal strength for vehicular ad hoc networks. In *IEEE Access*.



Rajiv Kapoor received M.E. and Ph.D. degree in Electronics and Communication Engineering from Delhi College of Engineering, Delhi University. Dr. Kapoor presently working as Professor in Electronics and Communication Engineering Department, AIACT&R (Govt. of NCT of Delhi). He has authored over 90 research papers in various renowned international journal and conferences. His primary research interests are machine learning, computer vision, signal and image processing.



Rashmi Gupta received M.E. and Ph.D. degree in Electronics and Communication Engineering from Delhi College of Engineering, Delhi University. Dr. Gupta held the position of Sr. Engineer in Calcom group of companies from 1991 to 1999, position of Lecturer in Electronics & Communication Engineering Department at Hindu Institute of Technology and Senior Lecturer in Maharaja Agrasen Intitute of Technology, Delhi from 1999 to 2003. Dr. Gupta presently working as Associate Professor in Electronics and Communication Engineering Department, AIACT&R (Govt. of NCT

of Delhi). She has authored over 40 research papers in various renowned international journal and conferences. Her primary research interests are machine learning, computer vision, signal and image processing.



Raghavendra Kumar is working as Assistant Professor in Computer Science and Engineering Department at L.N.C.T Group of College Jabalpur, M.P. India. He received B. Tech. in Computer Science and Engineering from SRM University Chennai (Tamil Nadu), India, M. Tech. in Computer Science and Engineering from KIIT University, Bhubaneswar, (Odisha) India and Ph.D. in Computer Science and Engineering from Jodhpur National University, Jodhpur (Rajasthan), India. He has published 86 research papers in international / National journal and conferences including IEEE, Springer and ACM as well as serve as session chair, Co-chair, Technical program Committee members in many international and national conferences and serve as guest editors in many special issues from reputed journals (Indexed By: Scopus, ESCI). He also received best paper award in IEEE Conference 2013 and Young Achiever Award-2016 by IEAE Association for his research work in the field of distributed database. His researches areas are Computer Networks, Data Mining, cloud computing and Secure Multiparty Computations, Theory of Computer Science and Design of Algorithms. He authored 12 computer science books in field of Data Mining, Robotics, Graph Theory, and Turing Machine by IGI Global Publication, USA, IOS Press Netherland, Lambert Publication, Scholar Press, Kataria Publication, Narosa, Edupedia Publication, S. Chand Publication and Laxmi Publication.



Le Hoang Son obtained the Ph.D. degree on Mathematics—Informatics at VNU University of Science, Vietnam National University (VNU). He has been promoted to Associate Professor in Information Technology since 2017. Currently, Dr. Son works as a researcher and Vice Director at the Center for High Performance Computing, VNU University of Science, Vietnam National University. His major field includes Artificial Intelligence, Data Mining, Soft Computing, Fuzzy Computing, Fuzzy Recommender Systems, and Geographic Information System. He is a member of International Association of Computer Science and Information Technology (IACSIT), Center for Applied Research in e-Health (eCARE), Vietnam Society for Applications of Mathematics (Vietsam). Dr. Son serves as Editorial Board of International Journal of Ambient Computing and Intelligence (IJACI, SCOPUS), Editorial Board of Vietnam Journal of Computer Science and Cybernetics (JCC), Associate Editor of International Journal of Engineering and Technology (IJET), Associate Editor of Neutrosophic Sets and Systems (NSS), and Associate Editor of Vietnam Research and Development on Information and Communication Technology (RD-ICT).



Sudan Jha holds the Ph.D. in CSE, working as professor in School of Computer Engineering, KIIT University, India with 18 years of experience in Academics, Administration as a Principal in four Colleges in India and Nepal, Industrial Research and Software Project Development and Management; As a team leader executed three government funded projects; Board of directors and Technical Advisor in Nepal's National Television (NTV); Chief IT Consultant in Nepal Telecom Authority – the regulating body of telecommunication of Nepal; published so far 31 International Journal transaction papers (SCOPUS Indexed-3, UGC Approved-8, Others-20); 60+ conference papers; 8 major software developed as Team Leader including Govt and non Govt. organization. Editorial board member in several Journals, International Conferences, Books. Provisional issuance is under process for his second Ph.D. in Computer Engineering which is already under the final verge. He has been an invitee for ROBOTRONICS USA twice, World Youth Festival Russia as a VIP Guest Speaker, Teacher's Education Club, Harvard University USA as an active audience. He has visited 14 countries and delivered talks in various topics related with Network Systems, Embedded Systems, Parallel & Distributed Systems, Operating System, Distributed Operating System, Programming Skills like Programming in C, Programming in C++, Object Oriented Software Engineering, Emerging Subjects like High Speed Downlink Packet Access Networks (HSDPA), Universal Mobile Universal Telecommunication Systems (UMTS), Cell throughput evaluation of HSDPA Networks, Fundamentals of Software Engineering, Data Mining and Warehousing, Wired and wireless networking, Software Engineering, Data Mining and Warehousing, UMTS based HSDPA. He also holds standing on consultant for the Nepal Telecommunication Authority, the regulatory board for Government of Nepal and member of board director of Nepal Television.

International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 6 Issue: 4 198 - 203

Optimization of Transfer Learning for Sign Language Recognition Targeting Mobile Platform

Dhruv Rathi
Department of Computer Engineering
Delhi Technological University
New Delhi, India

***Abstract*—**The target of this research is to experiment, iterate and recommend a system that is successful in recognition of American Sign Language(ASL). It is a challenging as well as an interesting problem that if solved will bring a leap in social and technological aspects alike. In this paper, we propose a real-time recognizer of ASL based on a mobile platform, so that it will have more accessibility and provides an ease of use. The technique implemented is Transfer Learning of new data of Hand gestures for alphabets in ASL to be modelled on various pre-trained high- end models and optimize the best model to run on a mobile platform considering the various limitations of the same during optimization. The data used consists of 27,455 images of 24 alphabets of ASL. The optimized model when ran over a memory-efficient mobile application, provides an accuracy of 95.03% of accurate recognition with an average recognition time of 2.42 seconds. This method ensures considerable discrimination in accuracy and recognition time than the previous research.

Keywords*-Sign Language; Inception; MobileNet; Transfer Learning; ASL; Mobile; Android*

*****

I. INTRODUCTION

The aimofdeveloping a real-time working system to accurately read the sign language and communicate with the signer in an automated sense, without having an expertise or even an introductory knowledge in that form of communication has been a topic of interest to the researchers of Computer Vision and Natural Language Processing, thus producing some noteworthy works which have had a great impact both in terms of algorithmic advancements and development of society. The problem of Sign Language Interpretation(SLI) is a form of the standard Action Classification Problem when seen from the aspect of Deep Learning. Majority of the previous works that have been able to set a benchmark in the field of SLI are such structured in their implementation that they are computationally expensive enough to be unsuitable for utilizing as a daily life tool limiting themselves through their dependency on hardware or are in limited to statically detecting and analyzing the gestures and sign languages, thus a robust and accurate system is needed.

In this work, the aim is to develop a novel real-time implemented system based on Mobile platforms that recognize and extracts sign actions of the American Sign Language(ASL) from live video frame to frame and construct sentences by monitoring the continuous action movements performed by the signer, i.e., the ASL will be converted to words in real time recording on an Android-based mobile phone. A visual framework for representation is provided that shows the corresponding likelihood of top 3 words and the time taken to for recognition.

The dataset used for the discussed research is Sign Language MNIST - Drop-In Replacement for MNIST for Hand Gesture Recognition Tasks [1]. A detailed discussion of the dataset is done in Section III(A).

To obtain a system that is accurate to the desired level and simultaneously holds the capability to run on a generic hardware deficient Android mobile phone, the system needs to be perfectly optimized to feed and satisfy itself with the low computational prowess. The best suitable approach to proceed with which will meet both the discussed parameters is the application of Transfer Learning. This is due to the possibility of high-optimality of the models and ease of training even with low amount of data as will be discussed in Section III(B).

Our proposed method uses Transfer Learning which makes the process and thus the system more robust with both large or small datasets. Training a powerful Deep Learning model from scratch can take days at a time and requires appreciable computational power and hardware, while Transfer Learning makes those models flexible and they easily adapt to new data. The obtained accuracy of recognition is 95.03%.

The remaining paper is catalogued as follows: Section II. Reviews the previous research done in the field. Section III: A) Delineates the dataset used B) Review and mathematics of Transfer Learning. Section IV: A) Discusses the experiment models and their structures B) Discusses the proposed experimentation processing C) Optimization of retrained model for Mobile Environment. Section V: A) Discusses and evaluates the results obtained from the experimentation B)

Method implementation and output C) Discusses the future work. Section VI: Mentions the references.

II. Previous Work

There are various approaches that have been proposed by researchers in the form of patents and research paper for the purpose of identification of gestures and signs pertaining to sign language.

G. Anantha Rao et. al. [2] discusses the results of the application of Deep Convolutional Neural Networks for Sign Language Recognition, with an accuracy of 92.88% recognition on a self-constructed dataset using OpenCV and Keras libraries.

Amrutha C U et. al. [3] in the paper "Improving language acquisition in sensory deficit individuals with mobile application" discusses construction of a mobile application that is a system of a speech-based transformer from text to Indian Sign Language (ISL) from a pre-built database of images of signs stored locally on the machine and are queried during runtime.

Sébastien Marcel et. al. [4] in the paper "Hand Posture Recognition in a Body-Face centred space" uses a neural network to interpret hand postures in an image. The dataset was self-constructed containing uniform and complex backgrounds with a recognition accuracy of 93.7% and 84.4% respectively.

III. Dataset Used And Transfer Leaning

A. *The Dataset Used*

The open dataset given at Kaggle called Sign Language MNIST - Drop-In Replacement for MNIST for Hand Gesture Recognition Task [1] which contains set of 28x28 images of all the alphabet, except J and Z, of the standard American Sign Language(ASL). The data contains a total of 27,455 cases. A sampled image set can be observed in Fig. (1)



Fig. (1)

The data in its raw form is provided as a pixel to pixel intensity [0-255] class-wise distributed XLS files. The data preprocessing steps included conversion of the mentioned data to image format using Python's open-source libraries Pandas, Numpy and others to obtain PNG format 28x28 grayscale images.

B. *What is Transfer Learning?*

Transfer Learning is an advanced technique of Deep Learning where a model developed for a task is used as a starting point for a model on a second context similar task. Training a Convolutional Neural Network(CNN) on ImageNet [give refer] dataset takes around 2-3 weeks across multiple GPUs. In the current research, two pre-trained models, Inception V3 model [5] and MobileNets [6] open-source models for comparison purposes. Both the mentioned models utilize CNN at their cores as discussed further.

In Transfer Learning technique, the last layers or some of the last layers of the pre-trained models, depending on the data being used are unfrozen and retrained on the new dataset to obtain required results. If only the last layer in the model is unfrozen, the resulting partial model then gives the probabilities of the outputs against each other rather than giving the most probable output, this altering will be much beneficial for sentence construction which will be discussed in the future work. If multiple last layers are unfrozen then a smaller Learning Rate is expected during retraining with the new dataset as this will help preserve the weights of the last few layers.

The MobileNets are small, low-latency, low-power models parametrized to meet the resources constraints of a mobile device. Google Research has 16 pre-trained open-source MobileNet models with varying parameters and accuracy competing for accuracy against time. MobileNets are mobile first Deepwise Separable Convolution Neural Networks, counting depthwise and pointwise convolutions as different layers, MobileNets has 28 layers and is trained on the ImageNet dataset.

The Inception V3 model that is used in the current research is a successor of Inception V1 model which is a variant of GoogleNet[7]with the additional features of Batch Normalization and the introduction of Factorization of larger convolutions into smaller convolutions. The main theory of models with Inception architecture is to find out how an optimal local sparse structure in a convolutional vision network can be approximated and covered by readily available dense components.

IV. Proposed Experimentation And Results

Here we present a method of experimentation that results into a mobile-based engine that successfully recognizes hand gestures and signs pertaining to the alphabets in the American Sign Language (ASL). The research is aimed specifically towards mobile-based system so that the system in real time implementation can be used as an everyday tool and does not suffer from mobility limitations. The first step in the experimentation is the construction of a Deep Learning model

that could successfully recognize the alphabet singularly after every requested frame, thus two models, 1) mobilenet_0.50_224 2) Inception V3 are trained using Transfer Learning and further optimized and further re-optimized to obtain a satisfactory model to run on mobile devices. Further, we will see implementation, experiments and results of both the models,

A. MobileNet_v1_0.50_224 model

The MobileNet_v1_0.50_224 model has 1.24 million parameters. The top layer of mobilenet_v1_0.50_224 takes as input a 1001-dimensional vector for each image, we train a softmax layer on top of this representation. Assume that the softmax layer contains N labels, this corresponds to learning N+1001*N model parameters corresponding to the learned biases and weights. When the process of Transfer Learning is initiated on the mentioned model, an analysis of all images is done and bottlenecks for each are calculated and stored, the bottleneck is referred to as the layer just before the final output layer that is responsible for the classification. This reasoning about why the training of the last layer alone still results in such a high accuracy is that the weights that are stored corresponding to each layer after forward and backward propagations during the long training durations are responsible for the classification of all the 1,000 objects in ImageNet will serve the same purpose with the new classes that we are aiming to classify with our provided data.

Our method as generalized showcases two variants of accuracy, Training accuracy, which is the measure of the percentage of images that were used in the current training batch and were correctly-labelled, another is Validation accuracy, which is the measure of precision on a randomly selected group of images from a varying set. It can be distinctly observed from the graphs in Fig. (4) and Fig. (6) that the validation accuracy fluctuates at irregular intervals, this is due to the reasoning that a random subset of the validation set is chosen for each validation accuracy measurement. Similar to as in Deep Learning training, the hyper-parameters are the ultimate game changers in Transfer Learning, several iterations were performed with varying hyperparameters, which in the present case includes, the Learning Rate (LR), which controls the magnitude of the update to the final layer during training.

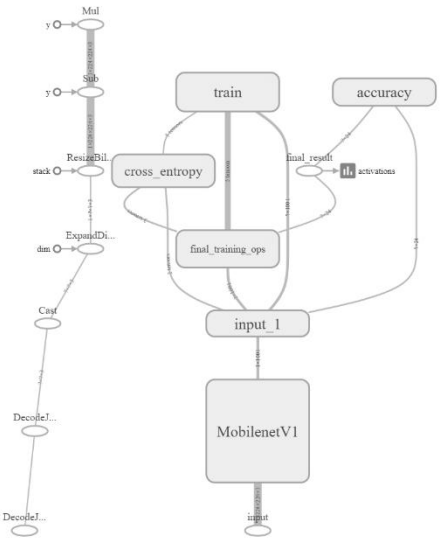


Fig. (2)

The MobileNet_v1_0.5_244 model is retrained on the mentioned dataset with an LR = 0.01 and number of training steps = 5000, with a training batch size of 100 images, from the dataset, the division of data is as follows, 10% data is kept for testing purposes, 10% data is kept for validation purposes and 80% data is used for training. On training the model on the later mentioned hardware, the training accuracy is 97.24% and the validation accuracy is 95.06%. Fig. (2) represents the graph obtained from Tensorboard, a visualization tool provided by Tensorflow [8], that provides an insight into the retrained MobileNet model and its architecture.

B. Inception V3 Model

The Inception V1 was developed as a variant of GoogleNet containing about 24 million parameters, with further versions V2 and V3, introducing Factorization of convolutions into smaller convolutions and addition of Batch Normalization to the fully connected layer respectively. The architecture of Inception model is a 299x299x3 input representing a field of 299 pixels and 3 channels in the image, combined with five convolutional layers, in addition to a handful of max-pooling operations, and successive stacks of Inception Modules, which are naively a set of different convolution filters and a max-pool filter combined undergoing filter concatenation. The output is a softmax layer at the end. The top layer of Inception V3 model takes as input a 2048-dimensional vector, we train a softmax layer on top of it. Assuming the softmax layer contains N labels, this corresponds to learning N + 2048*N parameters corresponding to learned weights and biases.

The hyper-parameters and accuracy metrics are similar as that of in the MobileNet case.

The Inception model is retrained on the mentioned dataset with an LR = 0.01 and number of training steps = 5000, with a training batch size of 100 images, from the dataset, the division of data is as follows, 10% data is kept for testing purposes, 10% data is kept for validation purposes and 80% data is used for training. On training the model on the later mentioned hardware, the training accuracy is 98.01% and the validation accuracy is 93.36%. Fig. (3) represents the graph obtained from Tensorboard, that provides an insight into the retrained Inception V3 model and its architecture.



Fig. (3)

C. *Optimization of a retrained model for Mobile Environment*

A modern smartphone can run about 10 Giga Floating Point Operations (GFLOPs) per second, thus the best expectant result from a 5 GFLOP model is two frames per second, this model and hardware dependency enables us to run the model even on limited RAM and high latency. Till now we have successfully retrained models that give us satisfactory accuracies for the purpose of Sign Language Recognition of alphabet in ASL, but when running any model on a mobile device, we might face three kinds of issues, 1) The Model and Binary Size; 2) App Speed and model loading speed; 3) Performance and Threading. Tensorflow [8] provides several benchmark and optimization tools that to counter the issues above.

The benchmark_model tool gives an estimate of the total number of floating point operations that are required to execute the graph, extrapolating this will lead us to get an idea of the minimum memory requirement for running the model considering in mind a satisfactory User Experience.

Model size is another potential part which can undergo high optimization. Transfer Learning models can be quite large in size given the huge training times and data involved, while a mobile phone has a limited memory that it has and that the operating system can allocate. Thus, the aim is to remove unusable and incompatible parts from the model, the important part of this is the number of constant parameters, let's assume M, stored as 32-bit float. A rough estimate of the size of the model is 4*N. Experimentation shows that only a small cutback in accuracy is found by replacing 32-bit by 8-bit per parameters, but the model size reduces by 75%.

Memory footprint is the most commendable technique to speed up the loading of models in the application. The idea is to load a file using memory mapping, rather than using the usual method of I/O APIs calling, in which an area of memory is allocated on the heap and then copying bytes from disk into it, the model is loaded directly into the memory, once the Operating System is notified through a required call, thus giving the advantage of 1) Increased speed in loading, 2) Reduces Paging, 3) The RAM budget dependency is thus removed from the whole system. Tensorflow handles this functionality with a function LoadMemoryMappedModel().

Thus we obtain an optimized model that is ready to be run on a mobile device, next step is the development of a JAVA based application for Android devices that incorporates at the least the following functionalities, 1) A camera interface that is used to capture the video frames during the act of performing the sign and gestures 2) An environment that is capable to process the developed model successfully, capture and record the results 3) An interface to indicate the results of the processing.

V. Discussion Of Results and its Significance

A. *Results and Comparison*

Upto this point the whole discussion revolves around experimentation with Inception V3 and MobileNet models, this is done with the aim to find the best model that will suit our needs, for the clarification of the same, we need to test the resulting models in real time environment and need a detailed visualization of the accuracy results of both,

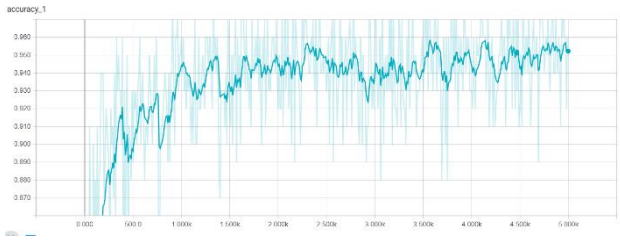


Fig. (4)

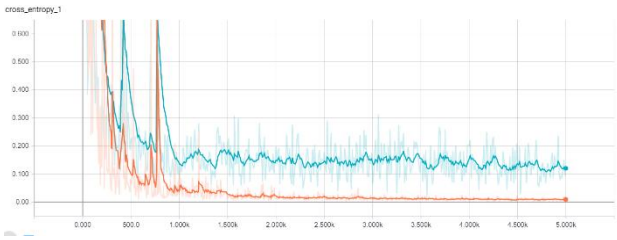


Fig. (5)

Fig. (4) is the Accuracy graph of Validation set of the retrained MobileNet model against training steps (5,000) and Fig. (5) represents the cross-entropy value of the retrained MobileNet model during training of the model, against training steps, colour coding of orange is indictive of Training data and Sky-blue that of Validation data, the final validation accuracy is 95.06%.

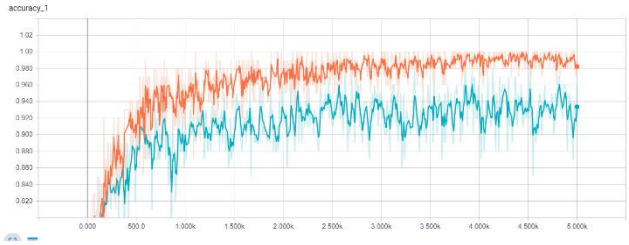


Fig. (6)

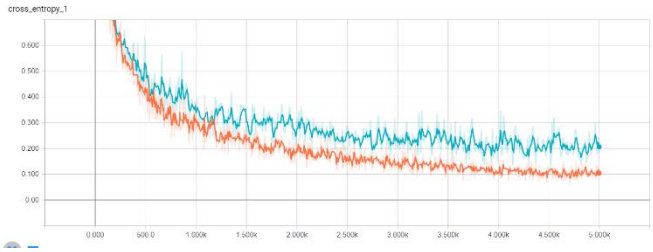


Fig. (7)

Fig. (6) is the Accuracy graph of the Training set and the Validation set of the retrained Inception V3 model against training steps (5,000) and Fig. (7) represents the cross-entropy value of the retrained Inception V3 model during training of the model, against training steps, colour coding of orange is indictive of Training data and Sky-blue that of Validation data, the final validation accuracy is 93.36%. The training accuracy, in this case, is 98.23% which indicates that the results are authentic and there is no instance of over-training.

S. No.	Model Name	Parameters (in million)	Time to retrain	Accuracy
1.	MobileNet	1.24	8.51 min.	95.06 %
2.	Inception_V3	24	12.31 min.	93.36%

Table (1)

From the output graphs, it is evident that MobileNet model outperforms Inception V3, although not by a very wide gap. But, upon implementation and testing the optimized models in a real time environment in a physical mobile device, the observation is that MobileNet outperforms Inception V3 by a very high measure in its accuracy of recognition of sign language alphabet, this statement can be attributed to different model architectures of the two models that will lead to learning of different parameters against the dataset provided and thus giving varied accuracies in recognition during the real time testing, as the images being processed contains numerous factors such as lighting conditions, varied backgrounds and others, which different models adapt to in their unique way.

The training of the models was done on the following configuration, NVIDIA GPU GeForce 940MX, 8GB RAM on Laptop hardware.

B. Real Time implementation and Output

An output of the real-time testing of the MobileNet modelled version of the application gives following outputs in recognition of the alphabet A, S and L. The boxes on the top of the screen in the results are indictive of the predicted probability of top three choices by the model.



A S L

Fig. (8)

Fig. (8) shows the ASL alphabet A, S and L with some noise in the background and achieve correct probabilities of 0.9988, 0.9681 and 0.9993 respectively. The other two probabilities for each of the three alphabets are separated by a considerably large difference so as to distinguishably declare the winner output.

C. Future Work

The discussed research has been successful in developing a system that can interpret American Sign Language alphabets in real time environment and thus can potentially act as a communication device between the signer and a non-signer and thus will help both technically and socially.

The researcher aims towards continuing the current ideation and system to include Natural Language Processing (NLP) and other techniques for sentence construction.

Also aimed at arethe construction ofproper datasets for other type of Sign Languages and inclusion of gestures and expanding the research by combining the previous point to develop a context-specific system for sentence construction that can ultimately act as a defined language interpreter.

VI. REFERENCES

[1] https://www.kaggle.com/datamunge/sign-language-mnist/, Sign Language MNIST, Kaggle, 2017.

[2] G.Anantha Rao, K.Syamala , P.V.V.Kishore, A.S.C.S.Sastry, Deep Convolutional Neural Networks for Sign Language Recognition, SPACES-2018, Dept. of ECE, K L Deemed to be UNIVERSITY, 2018.

[3] C. Amrutha C U,Nithya Davis, Samrutha K S, Shilpa N S, Job Chunkath, Improving language acquisition in sensory deficit individuals with, International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST, 2015.

[4] S. Marcel, Hand Posture Recognition in a Body-Face centred space, France Telecom CNET.

[5] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Rethinking the Inception Architecture for Computer Vision, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2818-2826., 2015.

[6] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, https://arxiv.org/abs/1704.04861, 2017.

[7] Christian Szegedy,Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, Going Deeper with Convolutions, https://arxiv.org/abs/1409.4842, 2014.

[8] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro,TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, https://arxiv.org/abs/1603.04467, 2016.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325145140

Performance Enhancement Analysis of Double Effect Vapor Absorption System Using Loop Heat Pipes

Article · January 2018

CITATIONS
0

READS
3

2 authors, including:



Ankit Dwivedi
Delhi Technological University
8 PUBLICATIONS **0** CITATIONS

SEE PROFILE

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

Doctoral Research View project

All content following this page was uploaded by Ankit Dwivedi on 15 May 2018.

The user has requested enhancement of the downloaded file.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

Performance Enhancement Analysis of Double Effect Vapor Absorption System Using Loop Heat Pipes

Ankit Dwivedi[1], R S Mishra[2]
[1, 2]*Dept. of Mechanical Engineering, Delhi Technological University, Delhi*

Abstract: ***A high temperature heat source cannot be used efficiently for a single effect vapor absorption refrigeration system (VARS). For the utilization of such high temperature source a double effect system can be brought into action. The COP I and COP II of a double effect VARS are generally higher than that of the single effect system. This research work implements some major modifications with the use of a Loop Heat Pipe (LHP) in the design of the double effect VARS to attain an enhanced performance. The simulations have found that the improved COP I and COP II are 2.09 and 0.4571 respectively. Also the percentage increase on the modified system has been recorded as 75.82% and 28.2 % respectively. Simulations also found that the average T_C can be maintained at 104.2ºC. The increase in the performance characteristics shows a positive gradient with the increase in the temperature of operation of LHP.***

I. INTRODUCTION

The fundamental of loop heat pipes(LHP) (Fig 1) is evaporation and condensation. On the heated side '1', the working liquid evaporates and on the cool side'3, 4, 5' it condenses. As every material has different properties (section A-A), it is required to choose the set of material properly. At the source the cool fluid is evaporated, the hot vapor stream is later moved to the sink through the vapor line'2', where the vapor condenses again and is transported back to the source through liquid line '5'. Highly efficient heat exchange is a direct result of the low heat resistance of vapors. This low heat resistance is because of little effective length of heat exchange through strong porous wick walls.

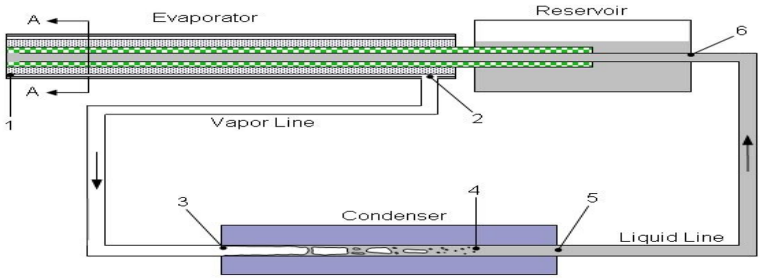


Fig 1: Cyclic process of a Loop Heat Pipe [35]

The LHPs have an inverted wick (Fig 2), and the vapors are located adjacent to the heated surface. These wicks are prepared by power metallurgy. The outer surface of the wick is in contact with the heated surface. Circumferential and axial grooves are required to generate flow channels for the vapors flow which can be machined in the wick, or to the evaporator body.

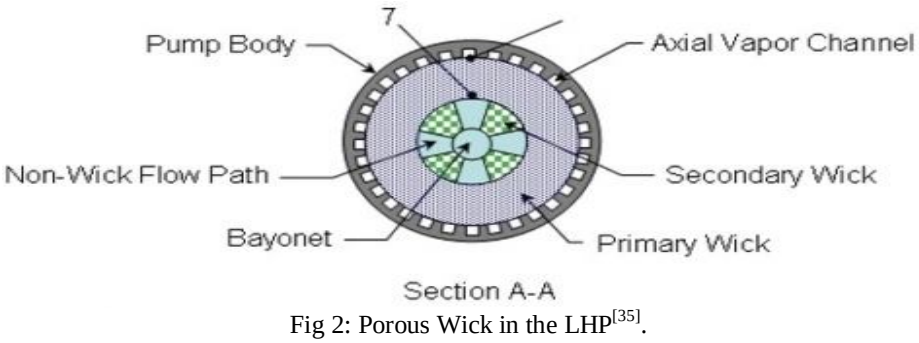


Fig 2: Porous Wick in the LHP[35].

©IJRASET (UGC Approved Journal): All Rights are Reserved

The single effect VARS is not well suited for the utilization of heat from a source with temperature higher than a certain temperature (The COP decreases with increase in temperature beyond a certain point). Hence the requirement of Double Effect VARS(Fig.3) is realised.

II. LITERATURE REVIEW

Fabian Korn et al. [2012]performed several vital experiments on heat pipes to establish it to be one of the most effective procedures to transport thermal energy from one point to another, mostly used for cooling[6].Sameer Khandekaret al.[2010]performed experiments on the global thermal performance modeling of Pulsating Heat Pipes (PHPs) requires local, spati-o-temporally coupled, flow and heat transfer information during the characteristic, self-sustained thermally driven oscillating Taylor bubble flow, under different operating conditions[7].JozefHužvár, Patrik Nemecet al. [2007]used heat pipe, observed its basic principles and operating limits. High temperature heat pipes were evaluated for use in energy conversion applications such as fuel cells, gas turbine re-combustors, and Stirling cycle heat sources, with the resurgence of space nuclear power, additional applications include reactor heat removal elements and radiator elements[8].R.Z. Wanget al.[2008]added heat pipes in adsorption water chiller or ice maker initials. His work showed that the adsorption refrigerators are very efficient [10].Pracha Yeunyongkul et al.[2009]aimedat experimentally investigating the application of a closed loop oscillating heat pipe (CLOHP) as the condenser for a vapor compression refrigeration system[14].R. Rajashree et al.[1990]went through a numerical analysis of an unsteady, viscous, laminar, incompressible, two dimensional heat and mass transfer, in the vapor gas region of gas loaded circular heat pipe [20]. Da-Wen Sun (1996) performed a detailed thermodynamic analysis of the properties of these binary fluids and expressed in polynomial equations. The performances of three cycles were compared. M.M. Talbi et al. (2000) carried out an exergy analysis on a single-effect absorption refrigeration cycle with lithium-bromide±water as the working Fuid pair. E. Kurem et al.(2001) analyzed the Absorption Heat Pump (AHP) and Absorption Heat Transformers (AHT) using ammonia-water and water-lithium bromide solutions. A fundamental AHP and AHT systems was described and explained the operating sequence. R.D. Misra et al. (2002) applied the therm-o-economic theory is to the economic optimization of a single effect water/LiBrvapor absorption refrigeration system for air-conditioning application. S.A. Adewusi et al (2004). studied the performance of single-stage and two-stage ammonia–water absorption refrigeration systems (ARSs).They calculated entropy generation of each component and the total entropy generation of all the system components as well as COP of the ARSs . S. Arivazhagan et al. (2006) investigated experimentally on the performance of a two-stage half effect vapor absorption cooling system .The prototype is designed for 1 kW cooling capacity using HFC based working fluids (R134a as refrigerant and DMAC as absorbent). Rabah Gomri et al. (2008) performed exergy analysis of double effect lithium bromide/water absorption refrigeration system. The system consisted of a second effect generator between the generator and condenser of the single effect absorption refrigeration system, including two solution heat exchangers between the absorber and the two generators. S.C. Kaushik et al. (2009) presented the energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption systems. They developed a computational model for the parametric investigation of the systems. Berhane H. Gebreslassie et al. (2010) performed an exergy analysis, which only considered the unavoidable exergy destruction, conducted for single, double, triple and half effect Water–Lithium bromide absorption cycles. Gulshan Sachdeva et al.(2014) performed anexergy analysis of VAR system using LiBr-H2O as working fluid with the modified Gouy-Stodola approach. Karl Ochsner (2008)et al. (2008) developed a new CO2-heat pipe with high-grade steel corrugated pipe system, which – contrary to other pipe systems permits raw length up to 100 m. They also described the establishment of the heat pump system in general. Research works have thoroughly studied thermodynamic and thermo-economic problems of the various VARS, namely half; single, double, triple effects etc. Some works have already done the exergetic analysis of the VAR systems. They work on the heat input, so these systems can be interconnected with other power generating systems, using the waste heat of the power development cycles. Also the heat is rejected from the condenser of VARS while the rich refrigerant is condensed. This heat can also be used to reduce the actual heat input in the generator. In the research work LHP can be used as the new component to be used for intra-cycle heat exchange and reduce the requirement of heat input for the cycle to operate. The LHP acts as a super conductor owing to its high heat transfer coefficients associated with the boiling and condensation. Also replacements of other bulky components by the LHP can be done, which would reduce the cost and size of the system.

III. SYSTEMS DESCRIPTION

Fig 3 shows a series flow double effect VARS working on Ammonia-Water (NH_3-H_2O) operating between 130ºC and 4ºC. It consists of 2 generators, a condenser, an evaporator, an absorber, a pump, 2 solution heat exchangers two solution reducing valves and a refrigerant expansion valves. This system has five temperature levels (temperature in LPG, temperature in HPG, evaporator

©IJRASET (UGC Approved Journal): All Rights are Reserved

temperature, condenser temperature, and absorber temperature) and three pressure levels (low pressure in the evaporator and absorber, medium pressure in the condenser and the low pressure generator, the high pressure in the high pressure generator). The strong solution is pumped to the HPG where it is heated to boil out the refrigerant vapor from the solution. The primary vapor, from the HPG moves to the LPG heating the medium concentration solution and then it's condensed. The heat of condensation of the primary vapor from the HPG is used in the LPG to get the secondary vapor. Thus, the total amount of liquid refrigerant leaving condenser is the sum of refrigerant originating from HPG and LPG. The refrigerant from the condenser expands into the evaporator where it extracts the heat of vaporization and cools the ambient. [7]

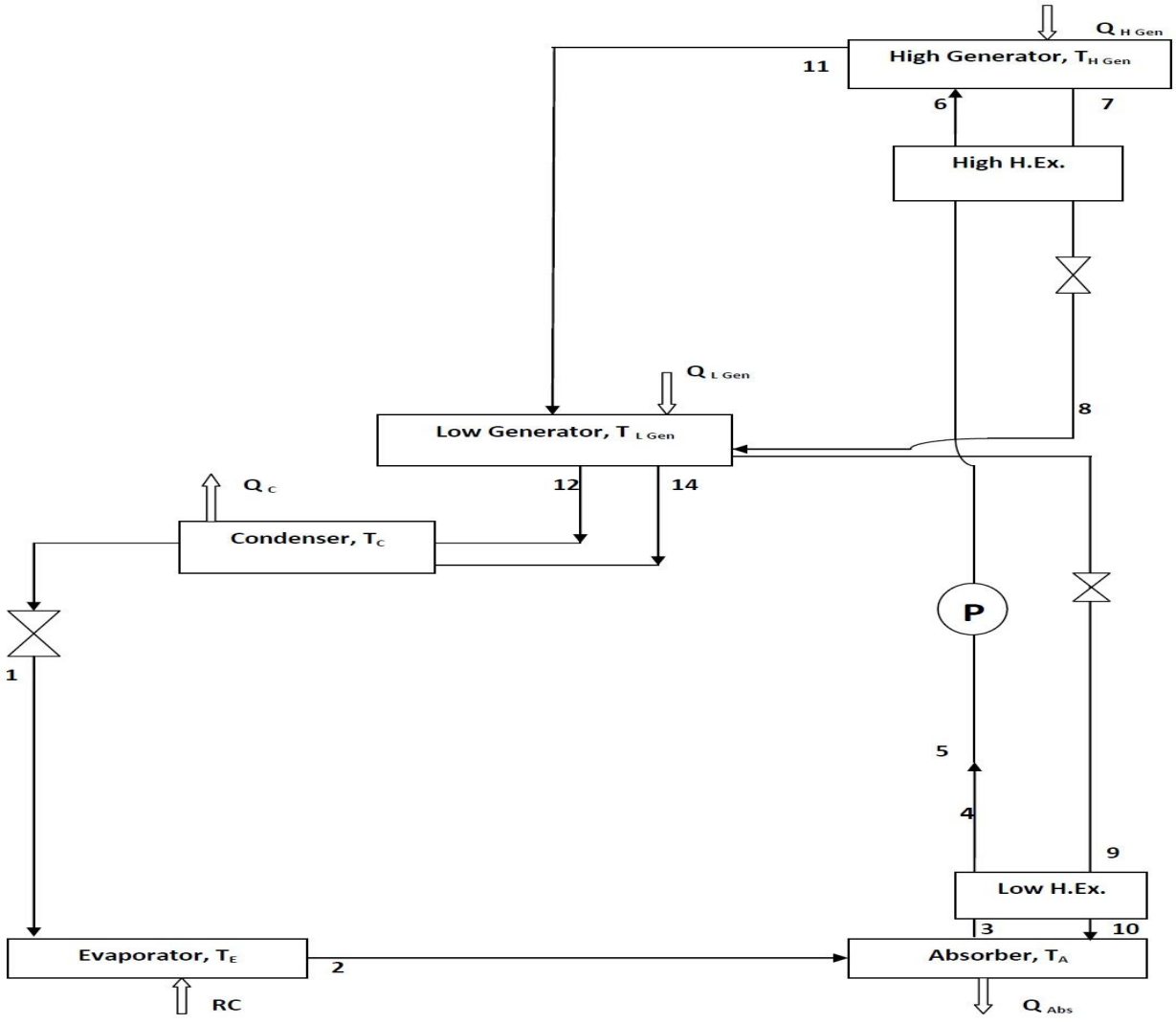


Fig 3: A Double Effect (Series)Vapor Absorption System

The proposed modifications in the double effect system have been showed in the fig 4. The main focus is on the reduction in the heat input in the HPG to improve the COP $_{I}$. Also the condenser is being replaced with the LHP which would be heating the mixture from the absorber using the heat which was going to be exhausted in the condenser. There are different materials that can be used to enhance the heat transferred to the working fluid of the LHP in the evaporator part. Effective wick area, pore radius, figure of merit etc affect the heat transfer factor of the heat pipe. [33] A better utilization of the waste heat will help in reducing the loss of exergy and in a way increasing the COP $_{II}$. Replacement of the condenser will also result in the total cost of the system, also it reduces the size.The LHP makes a cycle of heat exchange occur without any energy input.The table 1 consists of the terminology used in the analysis.

©IJRASET (UGC Approved Journal): All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

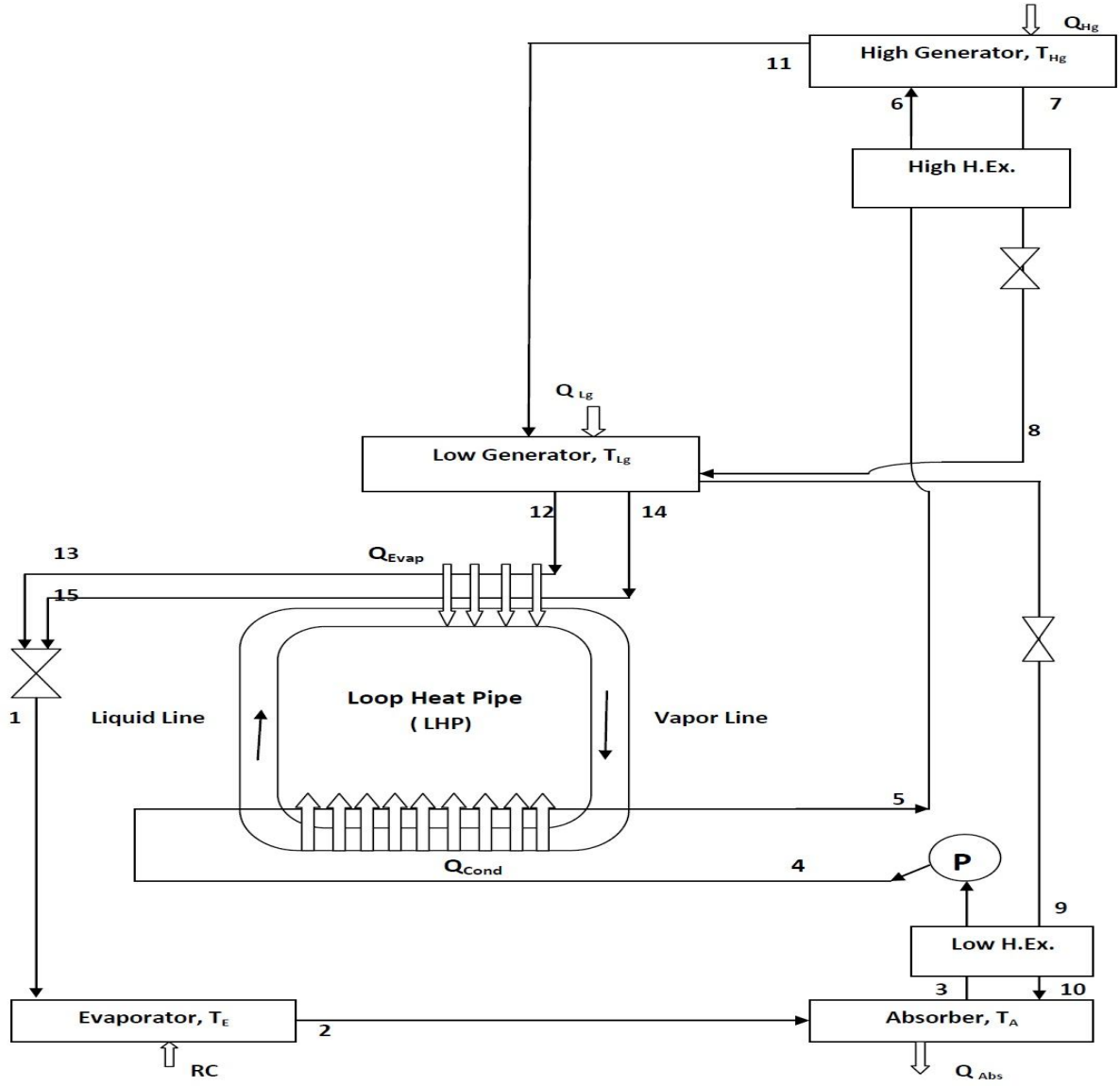


Fig.4: Modified Double Effect Vapor Absorption System

Table 1: Terms Used in Simulation

Terms	Abbreviations
Refrigeration Effect in kW	RE (kW)
Heat rejected in absorber in kW	Q_a(kW)
Heat supplied in generator in kW	Q_g (kW)
Heat rejected in condenser of LHP in kW	Q_{cond}(kW)
Heat absorbed in evaporator of LHP in kW	Q_{eva}(kW)
Absorber Temperature in°C	T_{La},T_{Ha}(°C)
Generator Temperature in °C	T_{Hg}, T_{LG} (°C)
LHP Condenser Temperature in °C	T_c (°C)
Evaporator Temperature in °C	T_E,T_e(°C)
Heat Rejected in Condenser in kW	Q_C (kW)

©IJRASET (UGC Approved Journal): All Rights are Reserved

First Law Coefficient of Performance	COP_{I}
Second Law Coefficient of Performance	COP_{II}
Heat Leaked from the LHP in kW	Q_{Leak} (kW)
Percentage Improvement in First Law Coefficient of Performance	$\%COP_{I\,imp}$
Percentage Improvement in Second Law Coefficient of Performance	$\%COP_{II\,imp}$
Improvement in First Law Coefficient of Performance	$COP_{I\,imp}$
Improvement in Second Law Coefficient of Performance	$COP_{II\,imp}$
Low Pressure Generator	LPG
High Pressure Generator	HPG

IV. RESULTS AND DICSUSSIONS

The Fig 5,6 and 7 show the variation of COP_{I} & COP_{II} with the heat being utilized in the LHP Q_{Cond}. Fig 5 shows the comparison COP_{I} of the modified system with the basic system.

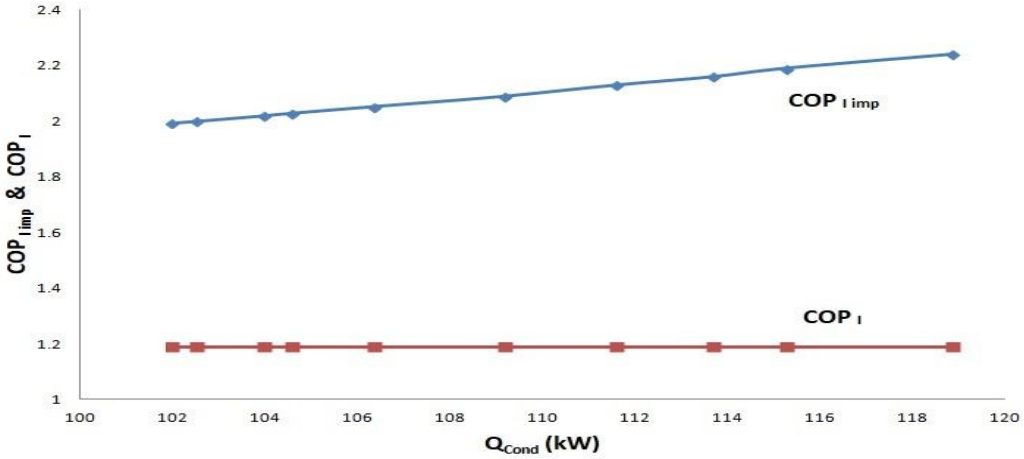


Fig 5: Comparison between COP_{I} and $COP_{I\,imp}$ plotted with Q_{Cond}

The average enhanced COP_{I} for the modified system is 2.09 where as the COP_{I} of the basic system is 1.189.

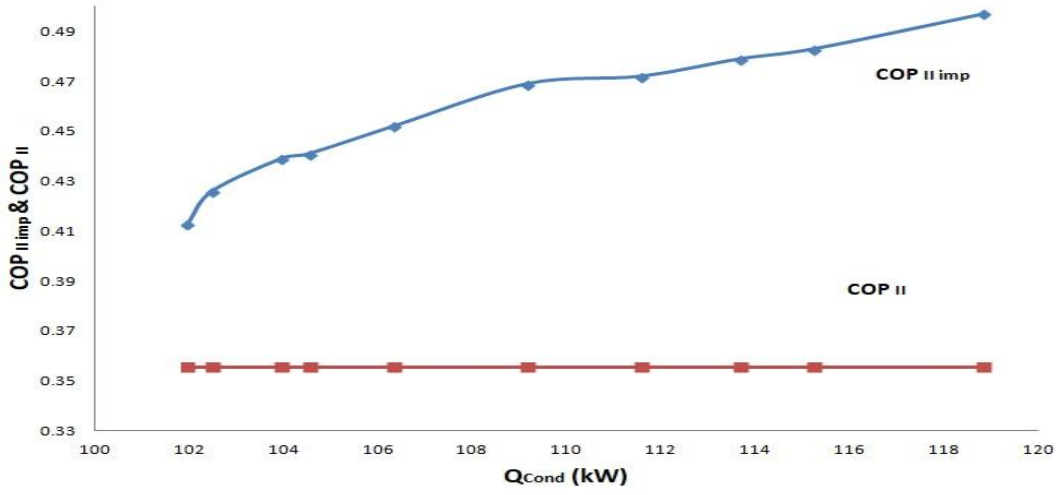


Fig 6: Comparison between COP_{II} and $COP_{II\,imp}$ plotted with Q_{Cond}

©IJRASET (UGC Approved Journal): All Rights are Reserved

The fig 6 describes the variation of COP $_{II}$ of the modified system with the basic system. The average COP $_{II}$ upon this modification is 0.4571, where the original COP $_{II}$is 0.3557.

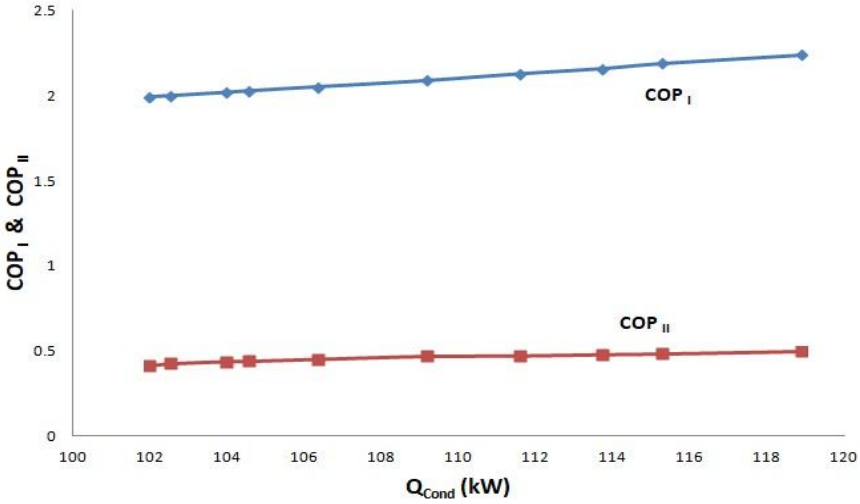


Fig 7: COP $_{I}$ and COP $_{II}$plotted with Q_{Cond}

The fig 7 compares the COP $_{I}$ and the COP $_{II}$ of the modified system with a variable Q_{Cond}. It can be seen that the increase in the performance parameters is gradual and consistent. More the heat utilized with the help of LHP, higher are the performances based on First law and Second law. It can be said that a modification inside the LHP will help in enhancing the performance. Fig 8, 9 and 10 help in analysing the %age improvement happening due to the modification of the system. Fig 8 shows the variation with the Q_{Cond}. The COP $_{II}$ has a sharper rise over the range when compared to the COP $_{I}$.

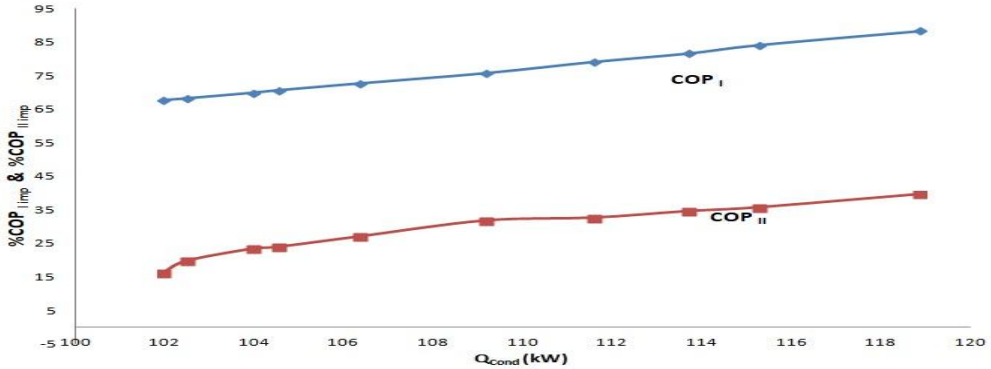


Fig8: Variation of %COP $_{I\,imp}$ and % COP $_{II\,imp}$ plotted with Q_{Cond}

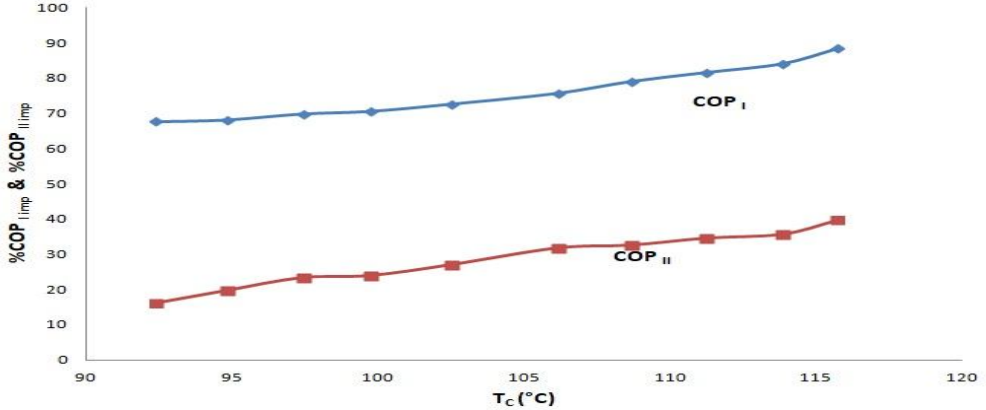


Fig9: Comparison of %age improvements in COP $_{I}$& COP $_{II}$ varying the T_C

©IJRASET (UGC Approved Journal): All Rights are Reserved

The fig 9 shows the %age rise in COP $_{I}$ and COP $_{II}$ with varying T_C. The COP I shows a increasing slope for the entire range while the COP II shows a decreasing slope. Further increase in the T_C may not result in the in the enhancement in the COP $_{II}$, as the COP $_{I}$ will show improvement for further increase in T_C. The average %age increase in COP I is 75.82.

Here fig 10 shows the variation of the above with a range of T_G. The trending is similar to that of fig 9. With the increase in the T_G the COP $_{I}$ is expected to increase (Fig 12). But with the rise in T_G also there will be a rise in Evaporator temperature of LHP along with the rise in T_C and decrease in the Q_{Leak}, hence the utilization of heat will increase. Along with this increases the COP $_{II}$ (Fig 11). The %age increase in COP $_{II}$ is 28.5.

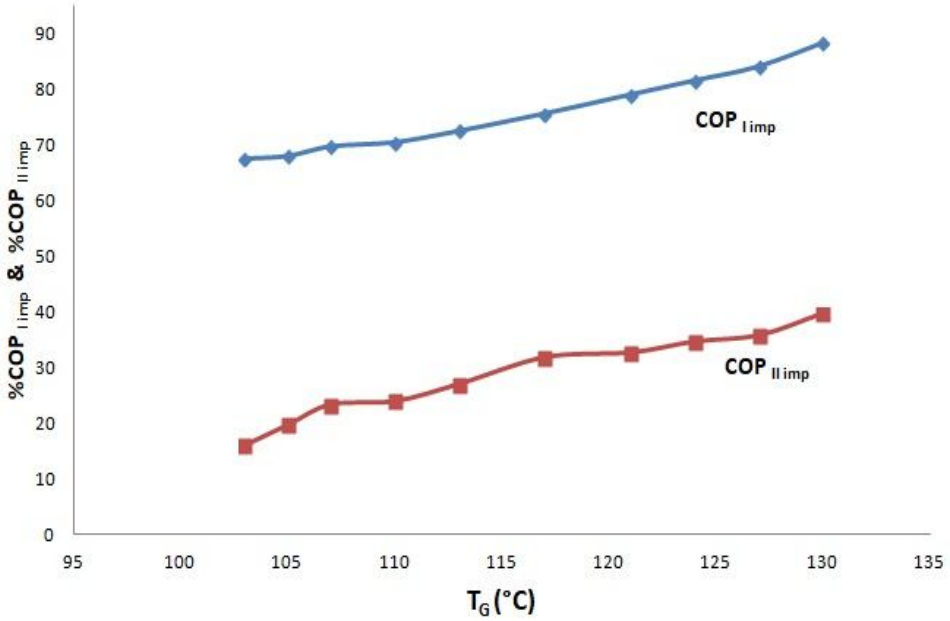


Fig 10: Comparison of %age improvements in COP $_{I}$& COP $_{II}$ varying the T_G

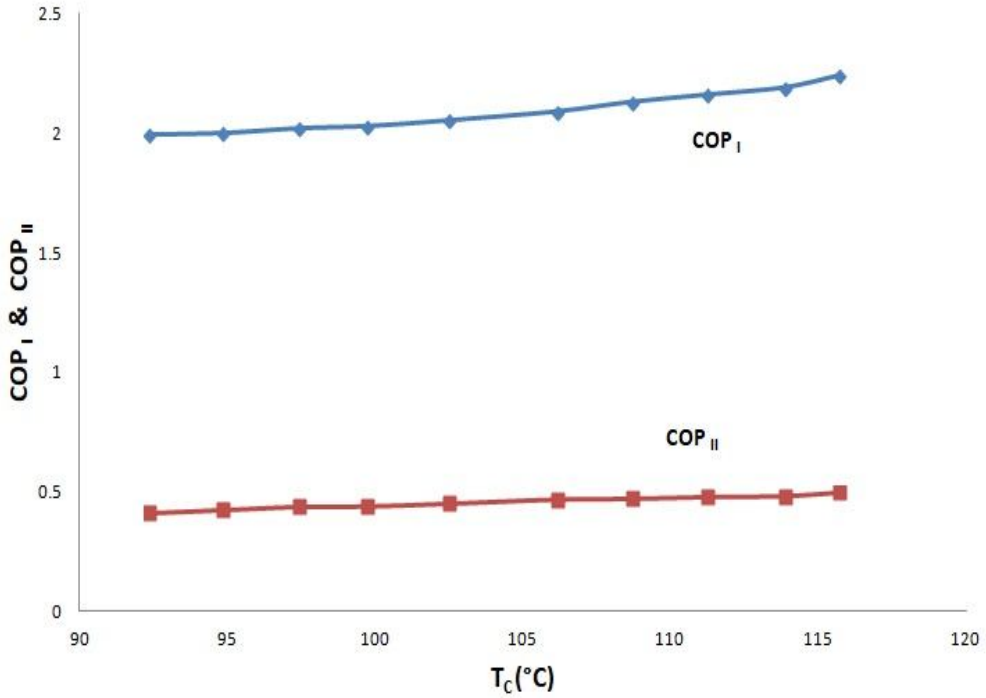


Fig 11: Comparison of COP $_{I}$& COP $_{II}$ varying the T_C

©IJRASET (UGC Approved Journal): All Rights are Reserved

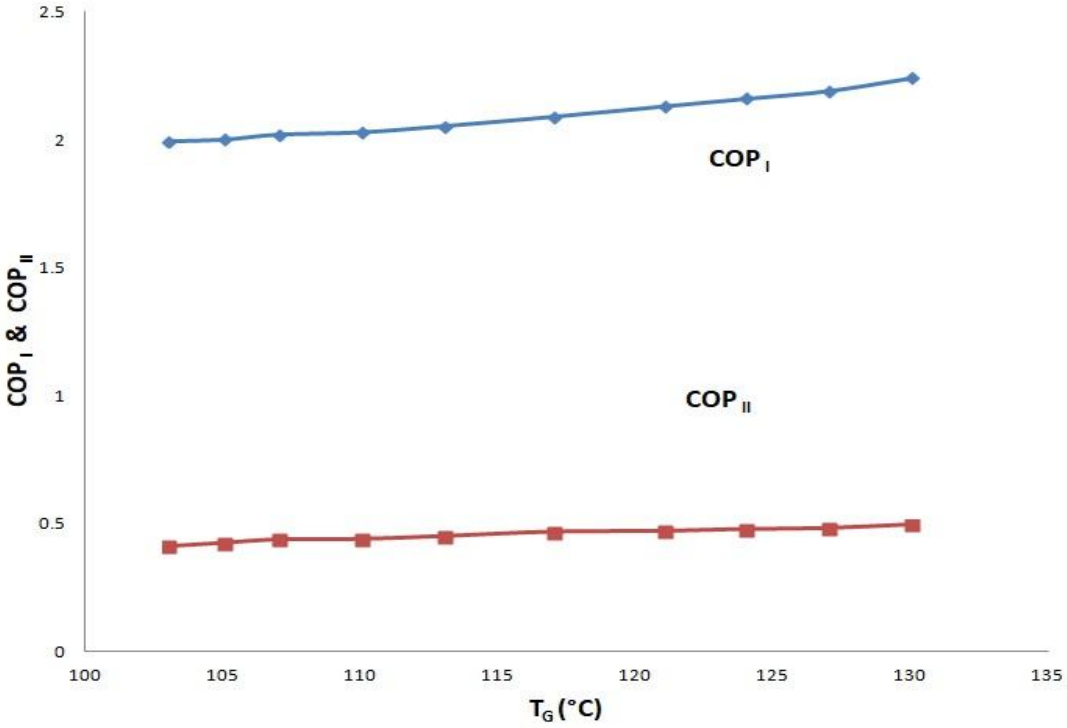


Fig 12: Comparison of COP $_{I}$& COP $_{II}$ varying the T_G

It has been observed that with increase in the T_G, COP $_{I}$, COP $_{II}$ and T_C all increase. Fig 13 and fig 14 show the variation of Q_{Cond} and Q_{Leak} with T_C and T_G respectively. In both the figures it can be easily noticed that the increase in the temperature provides desirable results such as reduction in the Q_{Leak} and increase in Q_{Cond}.

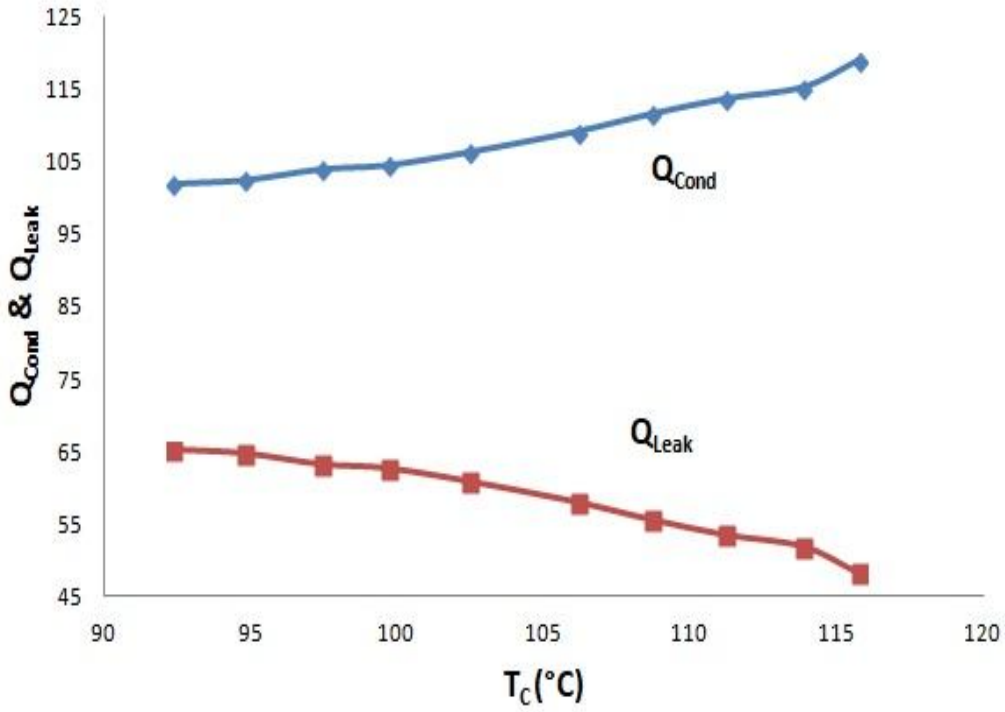


Fig 13: Comparison of Q_{Leak}&Q_{Cond}varying the T_C

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

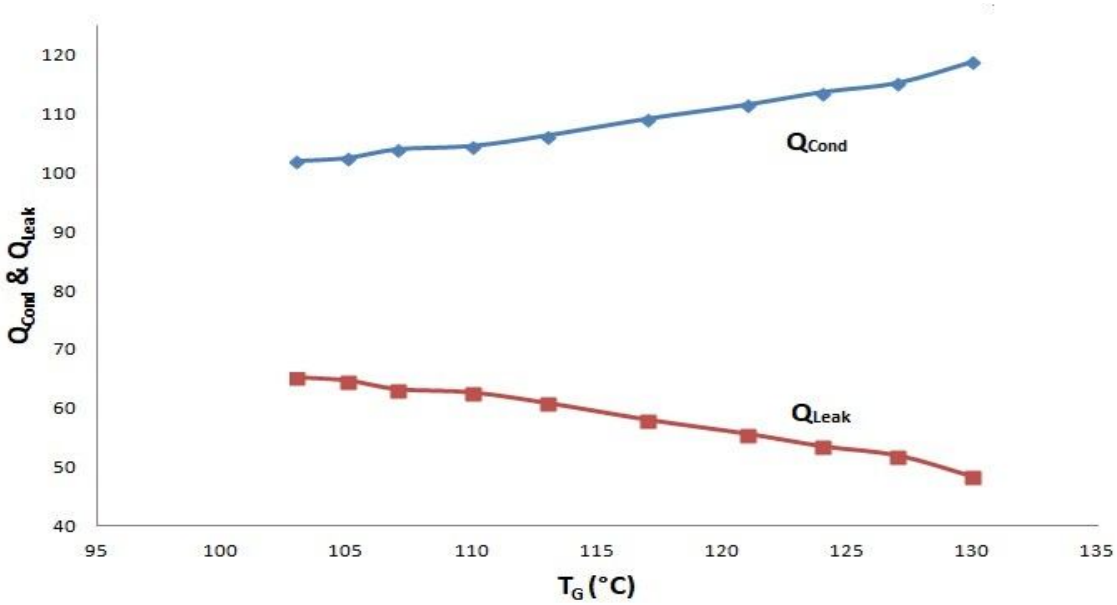


Fig 14: Comparison of Q_{Leak}&Q_{Cond} varying the T_G

The Q_{Cond} has a positive gradient in both the figures. Further increase in the temperature will increase the Q_{Cond} and hence the leakage of heat will decrease. The limit of the temperature is 130ºC; other methods have to be sorted out to increase the Q_{Cond}, such as changing the material and changing the arrangements of heat transfer surface of the LHP.

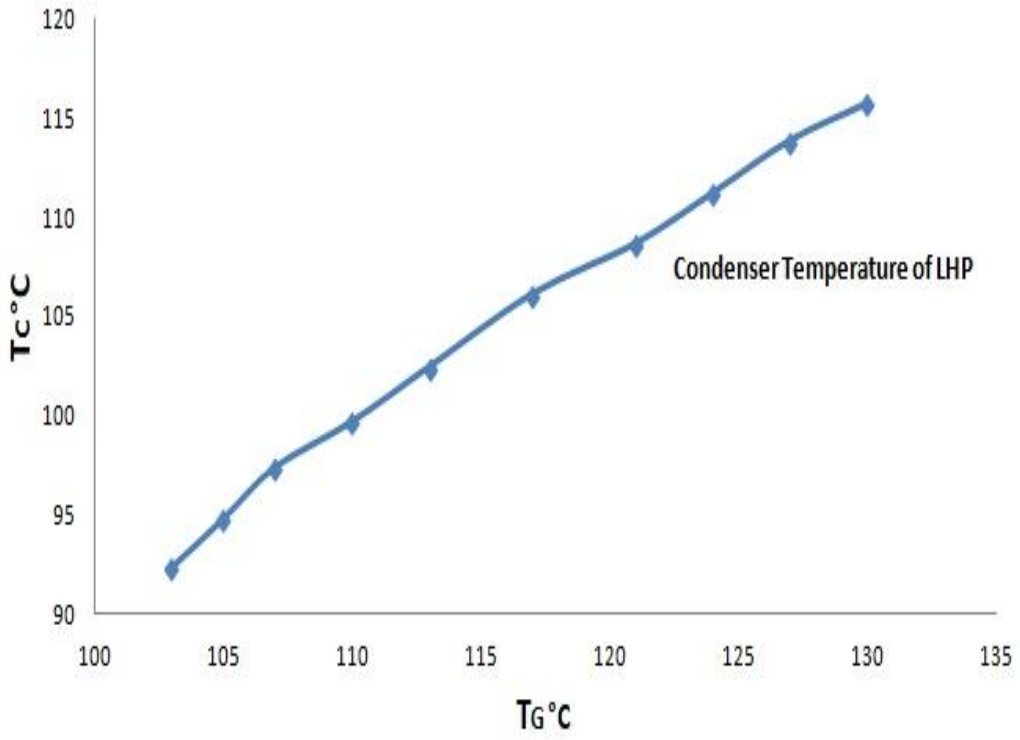


Fig 15: Variation of T_C with the T_G

Figure15 shows the increase in Condenser temperature of LHP. It helps in the availability of the heat to be utilized and the extent of heat transfer that can take place to the mixture. At the limit of the T_G the maximum temperature in the condenser of LHP is115ºC. Also on an average the T_C is found to be is 104.23 ºC.

©IJRASET (UGC Approved Journal): All Rights are Reserved

V. CONCLUSIONS

Following the results of the simulations, following conclusions can be made:

A. The average enhanced COP_I &COP_{II} are 2.09 and 0.4571 respectively.
B. The percentage increase in COP_I &COP_{II} can be observed to be 75.82 and 28.2 respectively.
C. The COP_{II} has a sharper rise over the range when compared to the COP_I for the range of Q_{Cond}.
D. The COP I shows a increasing slope for the entire range while the COP II shows a decreasing slope for the entire range of the T_G and T_C.
E. Increase in the temperature T_G provides the desirable results such as reduction in the Q_{Leak} and increase in Q_{Cond} and average T_C is found to be is 104.23 ºC .

REFERENCES

[1] Saeed. Sedigh , Hamid. Saffari ,"Thermodynamic analysis of single effect and half effect absorption refrigeration system" International Journal of Energy & Technology Vol.25 (2011) 1-9.
[2] S. Arivazhagan , R. Saravanan , S. Renganarayanan ," Experimental studies on HFC based two-stage half effect vapor absorption cooling system" Applied Thermal Engineering Vol. 26 (2006) 1455–1462.
[3] GulshanSachdeva, Ram Bilash," Thermodynamic Analysis of a Vapor Absorption System Using Modified Gouy-Stodola Equation" International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:8, No:12, 2014.
[4] I. Horuz," A comparison between Ammonia-water and Water-Lithium Bromide solutions in Vapor Absorption Refrigeration Systems" Int. Comm. Heat Mass Transfer, Vol. 25, No. 5, pp. 711-721, 1998.
[5] Abdul Khaliq, and Rajesh Kumar," Exergy analysis of double effect vapor absorption refrigeration system" Int. J. Energy Res. 2008; Vol.32:161–174.
[6] S.C. Kaushika, AkhileshArora," Energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption refrigeration systems" international journal of refrigeration Vol.32 (2009) 1247 – 1258.
[7] RabahGomri , RiadHakimi," Second law analysis of double effect vapor absorption cooler system" Energy Conversion and Management Vol.49 (2008) 3343–3348.
[8] S.A. Adewusi, Syed M. Zubair," Second law based thermodynamic analysis of ammonia–water absorption systems" Energy Conversion and Management Vol. 45 (2004) 2355–2369.
[9] RabahGomri," Second law comparison of single effect and double effect vapor absorption refrigeration systems" Energy Conversion and Management Vol. 50 (2009) 1279–1287.
[10] R.D. Misra, P.K. Sahoo, S. Sahoo, A. Gupta,"Thermoeconomic optimization of a single effect water/LiBrvapor absorption refrigeration system" International Journal of Refrigeration Vol.26 (2003) 158–169.
[11] M. Belghazi, A. Bontemps, C. Marvillet," Experimental study and modelling of heat transfer during condensation of pure fluid and binary mixture on a bundle of horizontal finned tubes" International Journal of Refrigeration Vol.26 (2003) 214–223.
[12] M.M. Talbi, B. Agnew," Exergy analysis: an absorption refrigerator using lithium bromide and water as the working Fuids" Applied Thermal Engineering Vol. 20 (2000) 619-630.
[13] E. Kurem," A comparison between Ammonia-water and Water-Lithium Bromide solutions in vapor absorption heat transformers "In. Comm. Heat Mass Transfer; Vol. 28, No. 3, pp. 421-438, 2001.
[14] Da-Wen Sun," Comparison of the performances and NH3-H20, NH3-LiNO3 and NH3-NaSCN Vapor Absorption Refrigeration Systems" Energy Convers. Mgmt Vol. 39, No. 5/6, pp. 357-368, 1998.
[15] Yu.F. Maydanik," Review Loop heat pipes" Applied Thermal Engineering Vol.25 (2005) 635–657.
[16] Randeep Singh, AliakbarAkbarzadeh , MasatakaMochizuk," Operational characteristics of a miniature loop heat pipe with flat evaporator" International Journal of Thermal Sciences Vol.47 (2008) 1504–1515.
[17] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, C.J. Chen," Performance study of a high efficient multifunction heat pipe type adsorption ice making system with novel mass and heat recovery processes" International Journal of Thermal Sciences Vol.46 (2007) 1267–1274
[18] Yuan-Ching Chiang , Wen-Cheng Kuo , Chia-Che Ho , Jen-JieChieh," Experimental study on thermal performances of heat pipes for air-conditioning systems influenced by magnetic nanofluids, external fields, and micro wicks" International Journal of Refrigeration Vol.43 (2014) 62 -70.
[19] T.X. Li, R.Z. Wang, L.W. Wang, Z.S. Lu," Experimental investigation of an innovative dual-mode chemisorption refrigeration system based on multifunction heat pipes" International Journal of Refrigeration 3 1 (2 0 0 8) 1 1 0 4 – 1 1 1 2.
[20] L. GarousiFarshi a,*, C.A. Infante Ferreira b, S.M.S. Mahmoudi a, M.A. Rosen," First and second law analysis of ammonia/salt absorption refrigeration systems" International Journal of Refrigeration Vol. 4 0 (2 0 1 4) 1 1 1-1 2 1.
[21] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, J.Y. Wu," Influence of mass recovery on the performance of a heat pipe type ammonia sorption refrigeration system using CaCl2/activated carbon as compound adsorbent" Applied Thermal Engineering Vol. 28 (2008) 1638–1646.
[22] Z.S. Lu , L.W. Wang, R.Z. Wang," Experimental analysis of an adsorption refrigerator with mass and heat-pipe heat recovery process" Energy Conversion and Management Vol. 53 (2012) 291–297.
[23] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences Vol.50 (2011) 592-598.
[24] Basant K. Agrawal*, Munawar N. Karimi," Thermodynamic performance assessment of a novel waste heat based triple effect refrigeration cycle" International Journal of Refrigeration Vol. 3 5 (2 0 1 2) 1 6 4 7-1 6 5 6.

©IJRASET (UGC Approved Journal): All Rights are Reserved

[25] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences 50 (2011) 592-598.
[26] T.S. Jadhav , M.M. Lele," Theoretical energy saving analysis of air conditioning system using heat pipe heat exchanger for Indian climatic zones" Engineering Science and Technology, an International Journal Vol. 18 (2015) 669-673.
[27] Chengchu Yan , Wenxing Shi , Xianting Li , Shengwei Wang," A seasonal cold storage system based on separate type heat pipe for sustainable building cooling" Renewable Energy Vol.85 (2016) 880-889.
[28] Matthias H. Buschmann," Nanofluids in thermosyphons and heat pipes: Overview of recent experiments and modelling approaches" International Journal of Thermal Sciences Vol.72 (2013) 1-17.
[29] P.D. Dunn, D.A. Reay, Heat Pipes, Pergamon Press, Oxford, 1993.
[30] D.Reay, Heat Pipes-Theory, Design and Applications, Butterworth-Heinemann,Oxford, Fifth edition 2006.
[31] Korn, F., "Heat Pipes and its Applications" Project Report 2008 MVK160 Heat and Mass Transport May 07, 2008, Lund, Sweden.
[32] Rajashree, R., Rao, K.S., "**A** Numerical Study of the Performance of Heat Pipe" Indian Journal of Pure and Applied Mathematics, 21 (1): 95-108, January 1990.
[33] C.P. Arora, Refrigeration and Air Conditioning, Tata Mcgraw-Hill Publishing Company Limited,Delhi, Third Edition, 2009.
[34] Ankit Dwivedi, R. S. Mishra," Thermodynamic Analysis of Heat Pipe Using Ammonia, Water and Ethanol with a View to Being Used in Refrigeration" ISSN 2347 - 3258 International Journal of Advance Research and Innovation, Volume 3, Issue 3 (2015) 498-502.
[35] Ankit Dwivedi, R. S. Mishra, Manjunath K,"Optimization of Vapor Absorption System Using Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET**),** Volume 5 Issue VIII, August 2017,634-639.
[36] https://teel.ucmerced.edu/introduction/loop-heat-pipe-simulation

©IJRASET (UGC Approved Journal): All Rights are Reserved

View publication stats

Journal of Non-Crystalline Solids 495 (2018) 85–94

Contents lists available at ScienceDirect

Journal of Non-Crystalline Solids

journal homepage: www.elsevier.com/locate/jnoncrysol

Photoluminescence investigations on Dy^{3+} ions doped Zinc Lead Tungsten Tellurite glasses for optoelectronic devices

Ritu Sharma, A.S. Rao*

Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi, 110 042, India

ARTICLE INFO

Keywords:
Glasses
Photoluminescence
Judd-Ofelt theory
Radiative properties
Emission cross-sections
W-LEDs

ABSTRACT

Dy^{3+} doped Zinc Lead Tungsten Tellurite (ZnPbWTe) glasses have been synthesized by conventional melt quench method and their luminescence behavior was analyzed using optical absorption, excitation, photoluminescence (PL) and PL decay spectral studies. The Judd-Ofelt (J-O) intensity parameters evaluated from the absorption spectra are used to determine various radiative parameters such as transition probability (A_R), branching ratio (β_R) and radiative lifetimes (τ_R) for the prominent emission transitions of Dy^{3+} ions in ZnPbWTe glasses. Under 384 nm excitation, PL spectra recorded for the titled glasses, exhibit two intense peaks; one in yellow at 575 nm and the other one in blue at 483 nm region. Among these two, yellow band ($^4F_{9/2} \rightarrow {}^6H_{13/2}$) is relatively more intense than the blue one ($^4F_{9/2} \rightarrow {}^6H_{15/2}$). The intensity of PL spectra increases up to 1 mol% of Dy^{3+} ions in ZnPbWTe glasses and beyond concentration quenching is observed. Branching ratios (β_R) and emission cross-sections (σ_{se}) were estimated for $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transition to understand the potential utility of these glasses for laser action in visible region. The CIE chromaticity coordinates, color correlated temperature (CCT), color purity and Y/B ratios were also estimated to understand the suitability of these glasses for w-LEDs. The luminescence quantum efficiency evaluated for the titled glasses showing highest value for 1 mol% of Dy^{3+} ions present in the as prepared glasses. From the measured emission cross-sections, quantum efficiency and CIE coordinates, it was observed that 1 mol% of Dy^{3+} ions in ZnPbWTe glasses are aptly suitable for optoelectronic devices such as lasers and w-LEDs.

1. Introduction

Recently, Rare Earth (RE) doped glassy and crystalline materials have attracted a great deal of attention because of their applications in diversified fields such as modern lighting technology, display, telecommunications, photovoltaics, optical fibers, solid state lasers and mid-infrared lasers etc. [1–5]. Generally, RE doped glasses are superior to crystalline materials due to their homogeneous light emitting capacity, simple manufacturing procedures, low production cost and better thermal stability. Currently, researchers working in the field of optoelectronic devices are paying more attention in developing visible optical devices suitable for white light emitting diode (w-LED) [6]. Further, w-LEDs have received much attention owing to their excellent properties such as environment friendliness, energy saving, small volume, long life span and high reliability. In order to improve benefits related to environment and reducing consumption of global energy as well as fossil fuels, w-LEDs are suitable replacement for incandescent and fluorescent lamps [7–9].

Tellurite based heavy metal oxide glasses are very interesting to study as they are having special characteristic features such as low phonon energies, high refractive index, high dielectric constant, low melting temperature and good RE ion solubility [10]. It is well known that the materials containing tungsten oxide (WO_3) possess electrochromic as well as photo-chromic properties and are useful in smart windows, display devices and sensors [11,12]. Further, WO_3 suppresses non-radiative losses and enhances quantum efficiency [13]. To obtain more efficient laser system, heavy metal compound such as lead oxide is added to the conventional glasses as its presence can improve the effective fluorescence [14]. The lead oxide (PbO) is having ability to form stable glasses over a wide range of concentrations due to its large mass, low field strength, high polarizability and thereby acting as glass former/modifier. Further remarkable characteristic of PbO as glass modifier in PbO_6 structural unit and glass former in PbO_4 structural unit, make glasses moisture resistant [15]. Furthermore, lead oxide based glasses possess high refractive indices, wide infrared transmittance, good thermal and chemical stabilities and high spontaneous emission probabilities [16–18]. Among the glass components, ZnO based materials have received significant consideration due to its low

* Corresponding author.
E-mail address: drsrallam@gmail.com (A.S. Rao).

https://doi.org/10.1016/j.jnoncrysol.2018.04.056
Received 2 February 2018; Received in revised form 24 April 2018; Accepted 27 April 2018
0022-3093/ © 2018 Elsevier B.V. All rights reserved.

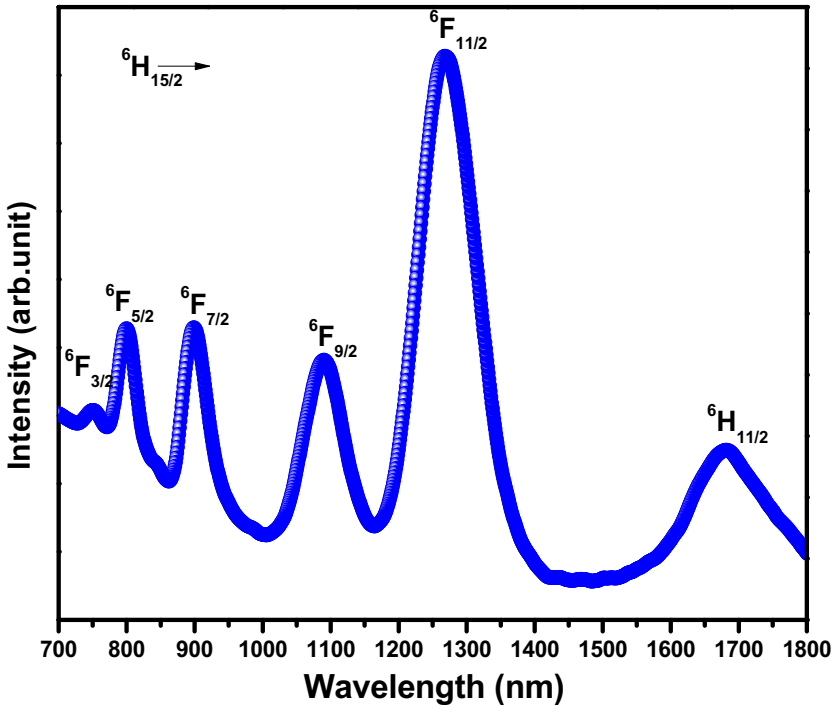


Fig. 1. Absorption spectrum of 1 mol% Dy^{3+} ions (glass B) in ZnPbWTe glass.

cost, direct wide band gap, intrinsic emitting property, large excitation binding energy, nontoxic and hygroscopic nature [19]. ZnO in the host glass exhibits dual nature as glass former as well as network modifier. Further, ZnO containing glasses are having significant importance due to its thermally stable and sublime nature. The aforementioned various scientific patronages possessed by the chemical constituents such as ZnO, PbO, WO_3 and TeO_2 prompted us to prepare a glassy system namely Zinc Lead Tungsten Tellurite (ZnPbWTe) glasses doped with different RE ions to study the effect of glass composition on the luminescence behavior of the doped RE ions.

Among the RE doped glasses, Dy^{3+} activated luminescent materials have attracted much attention and revolutionized the modern lighting and display fields, as it is capable of emitting ample wavelengths based on their 4f-4f transitions [20]. It is well known that the visible luminescence of trivalent dysprosium (Dy^{3+}) possess mainly two bands in the visible region corresponding to the transition $^4F_{9/2}\rightarrow{}^6H_{15/2}$ (blue) and $^4F_{9/2}\rightarrow{}^6H_{13/2}$ (yellow). It is important to optimize the local environment to obtain white light from Dy^{3+} doped glasses at suitable yellow to blue (Y/B) intensity ratio [21]. By changing the content of Dy^{3+} and/or composition of the host material white light emission at a modulated yellow to blue (Y/B) intensity ratio is possible [22]. The lasing action of Dy^{3+} ions in visible region also finds several applications in diversified fields of science and technology such as commercial display sensors and fiber optic amplification [23]. In the present work ZnPbWTe glasses are prepared by varying Dy^{3+} ion concentration and investigated thoroughly using the absorption, emission and lifetime measurements to select the efficient host material for solid state laser and w-LED applications.

2. Experimental

2.1. Synthesis

Dy^{3+} doped ZnPbWTe glasses were prepared by conventional melt quenching technique with the following composition.

$(60 - x)TeO_2 + 20WO_3 + 15PbO + 5ZnO + xDy_2O_3$ (where x = 0.5 mol% for glass A, x = 1.0 mol% for glass B, x = 1.5 mol% for glass C, x = 2.0 mol% for glass D and x = 2.5 mol% for glass E). Analar grade chemicals such as TeO_2, WO_3, PbO, ZnO and Dy_2O_3 were taken as the starting materials. The raw materials were thoroughly mixed in an agate mortar until a smooth and homogeneous mixture was obtained. The mixture thus obtained is poured into a silica crucible and heated at 830 °C in an electric furnace for about 55 min to obtain bubble free homogeneous melt. Such melts were then quenched suddenly on a pre-heated brass mould and pressed quickly by another brass mould. The prepared glass samples were subsequently annealed at 250 °C for 1 h to remove thermal strains/cracks and then polished to do spectral measurement.

2.2. Characterization

Density and refractive index are the two important parameters to study the structural modification of glass network. For all the glasses under investigation densities were measured by employing the Archimedes method using water as an immersion liquid. The refractive indices were measured by Brewster's angle method using He-Ne laser (650 nm line) as the source. Optical absorption spectra have been recorded for all the ZnPbWTe glasses in NIR region using Perkin Elmer UV–vis-NIR spectrometer Lambda 950 (USA). The excitation and emission spectra for all the glasses were recorded using RF-5301PC Spectrofluorophotometer at room temperature. Fluorescence decay spectral recordings were carried out using Edinburgh FLSP900 fluorescence spectrometer having 0.1 nm spectral resolution. All the figures are plotted with the help of origin 9.2 software.

3. Result and discussion

3.1. Absorption spectral measurements (hypersensitive transition and J-O intensity parameters)

Fig. 1 shows the absorption spectrum of glass B recorded at room temperature in the wavelength range 700–1800 nm. The absorption spectra for all the as prepared glasses are alike in absorption features with slight variation in intensity. Hence absorption spectra of the

Table 1
Experimental ($f_{exp} \times 10^{-6}$), calculated ($f_{cal} \times 10^{-6}$) oscillator strengths and r.m.s deviation for Dy^{3+} ions in ZnPbWTe glasses.

Transitions from $^6H_{15/2}$→	Glass A		Glass B		Glass C		Glass D		Glass E	
	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}
$^6F_{3/2}$	0.31	0.42	0.46	0.49	0.65	0.60	0.44	0.50	0.29	0.42
$^6F_{5/2}$	2.81	2.25	3.43	2.62	4.74	3.22	3.89	2.68	3.05	2.28
$^6F_{7/2}$	4.27	4.57	5.07	5.52	5.63	6.49	4.92	5.59	4.43	4.85
$^6F_{9/2}$	5.02	4.96	6.54	6.45	7.09	6.92	6.54	6.40	5.86	5.77
$^6F_{11/2}$	9.93	10.1	10.38	10.5	11.52	11.6	10.80	10.75	10.02	10.2
$^6H_{11/2}$	5.89	5.89	6.09	6.09	6.78	6.78	6.32	6.32	5.98	5.98
$\delta_{rms}(\times 10^{-6})$	± 1.44		± 1.41		± 1.53		± 0.62		± 1.50	
$\delta_{rms}^{**}(\times 10^{-6})$	± 0.290		± 0.42		± 0.78		± 0.62		± 0.39	

δ_{rms} is including hypersensitive transition ($^6H_{15/2} \rightarrow {}^6F_{11/2}$).
δ_{rms}^{**} is excluding hypersensitive transition ($^6H_{15/2} \rightarrow {}^6F_{11/2}$).

remaining glasses are not shown. The absorption spectrum consists of six distinct absorption bands at 750, 800, 899, 1090, 1269 and 1680 nm which correspond to the $^6H_{15/2} \rightarrow {}^6F_{3/2}$, $^6F_{5/2}$, $^6F_{7/2}$, $^6F_{9/2}$, $^6F_{11/2}$ and $^6H_{11/2}$ transitions respectively. The assignments of the absorption transitions have been done according to Carnall et al. [24] and identification of bands is shown in Fig. 1. To evaluate the energies of the absorption bands in terms of experimental oscillator strength (f_{exp}), area under absorption curve is used by employing the equation given in literature [25] and the values are tabulated in Table 1. From the estimated (f_{exp}) values the J-O intensity parameters (Ω_2, Ω_4, Ω_6) and calculated oscillator strengths (f_{cal}) are evaluated by least square fit method using the J-O theory [26,27]. Root mean square deviation (δ_{rms}) was used to evaluate the quality of fit between f_{exp} and f_{cal} strength by using the following expression.

$$\delta_{rms} = \sqrt{\sum (f_{exp} - f_{cal})^2 / N}$$

here, N represents the total number of energy levels present in the fitting procedure. The f_{exp} and f_{cal} parameters along with δ_{rms} including and excluding hypersensitive transition ($^6H_{15/2} \rightarrow {}^6F_{11/2}$) are presented in Table 1. The goodness of fit between f_{exp} and f_{cal} as well as the accuracy of J-O intensity parameters Ω_λ (λ = 2, 4, 6) are established by considerably small δ_{rms} values, when hypersensitive transition is excluded. For the titled glasses the J-O intensity parameters are following the same trend ($\Omega_2 > \Omega_6 > \Omega_4$). Table 2 represents the evaluated J-O intensity parameters for the as prepared glasses under investigation along with comparison of J-O intensity parameters in various other glass matrices [21,23,25,28–33]. The magnitudes of Ω_λ (λ = 2, 4, 6) suggest that the observed trend of the intensity parameters is due to the ligand environment around the RE ion. According to Jorgensen and Reisfeld [34], the factors which influence the magnitude of the Ω_2 parameter are covalent nature of metal–ligand bond and asymmetry of ion sites in the neighbourhood of RE ion along with local structure and composition of the glass. The structure dependent parameters Ω_4 and Ω_6 are related to the rigidity and viscosity of the medium in which ions are situated. Higher value of Ω_2 parameter in Table 2 elucidates the covalent nature of RE ion (Dy^{3+}) with ligand and lower symmetry of RE ions in the host glasses. It is well known from literature that hypersensitivity is described by the magnitude of intensity parameter Ω_2 because hypersensitive transitions are associated with the large values of oscillator strengths as well as reduced matrix element $\|U^2\|$. The degree of hypersensitivity can be investigated by studying the variation of Ω_2 in different environments [26]. In any host hypersensitive transition possesses more intensity than other transitions and follows the selection rules $\Delta S = 0, |\Delta L| \leq 2$, $|\Delta J| \leq 2$ [35].From Tables 1 and 2 it can be concluded that in the titled glasses, the transition $^6H_{15/2} \rightarrow {}^6F_{11/2}$ is hypersensitive with higher magnitude of f_{exp} which depends on the magnitude of Ω_2. The variation of oscillator strength for hypersensitive transition ($^6F_{11/2}$) and ($\Sigma\Omega_\lambda$) with Dy^{3+} ions concentration is represented in Fig. 2. It is clear from Fig. 2 that, both the parameters are varying in the same manner with Dy^{3+} ions concentration except intensity.

3.2. Excitation and photoluminescence spectral measurements

In order to predict the luminescence properties of the as prepared glasses, it is important to select suitable excitation wavelength. Therefore the excitation spectrum of 1 mol% of Dy^{3+} ions in ZnPbWTe glass is recorded by monitoring the emission at 575 nm and is shown in Fig. 3. Four excitation bands are observed in Fig. 3 at 349, 363, 384 and 424 nm corresponding to the transition $^6H_{15/2} \rightarrow {}^6P_{7/2}$, $^6P_{5/2}$, $^4I_{13/2}$, $^4G_{11/2}$ respectively. It is well-known that the wavelength corresponding to the prominent excitation band can give intense emission. Since the excitation band at 384 nm ($^6H_{15/2} \rightarrow {}^4I_{13/2}$) is found to be more intense, hence all the ZnPbWTe glasses have been excited with 384 nm

Table 2
Judd-Ofelt Parameters ($\Omega_\lambda \times 10^{-20}$ cm^2) of Dy^{3+} ions in ZnPbWTe glass along with various reported hosts.

Glass system	Ω_2	Ω_4	Ω_6	Trend	References
Glass A	3.84 ± 0.01	1.57 ± 0.01	2.81 ± 0.01	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
Glass B	3.56 ± 0.01	2.44 ± 0.01	3.17 ± 0.01	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
Glass C	3.71 ± 0.02	1.99 ± 0.04	3.26 ± 0.09	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
Glass D	3.67 ± 0.01	2.27 ± 0.01	3.19 ± 0.01	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
Glass E	3.78 ± 0.01	2.31 ± 0.01	2.78 ± 0.01	$\Omega_2 > \Omega_6 > \Omega_4$	Present work
PbFPDy10	7.21	1.59	2.20	$\Omega_2 > \Omega_6 > \Omega_4$	[21]
ZnAlBiBDy	12.03	2.33	2.74	$\Omega_2 > \Omega_6 > \Omega_4$	[23]
L5FBDy	9.01	2.56	3.52	$\Omega_2 > \Omega_4 = \Omega_6$	[25]
$LiNbO_3$	5.42	1.14	2.51	$\Omega_2 > \Omega_6 > \Omega_4$	[28]
LSG	19.08	1.88	6.04	$\Omega_2 > \Omega_6 > \Omega_4$	[29]
Fluorosilicate	8.14	2.04	2.26	$\Omega_2 > \Omega_6 > \Omega_4$	[30]
D1LBE	6.45	2.02	3.87	$\Omega_2 > \Omega_6 > \Omega_4$	[31]
PYBDy10	7.75	2.31	2.70	$\Omega_2 > \Omega_6 > \Omega_4$	[32]
Lead Borate	4.90	0.94	2.07	$\Omega_2 > \Omega_6 > \Omega_4$	[33]

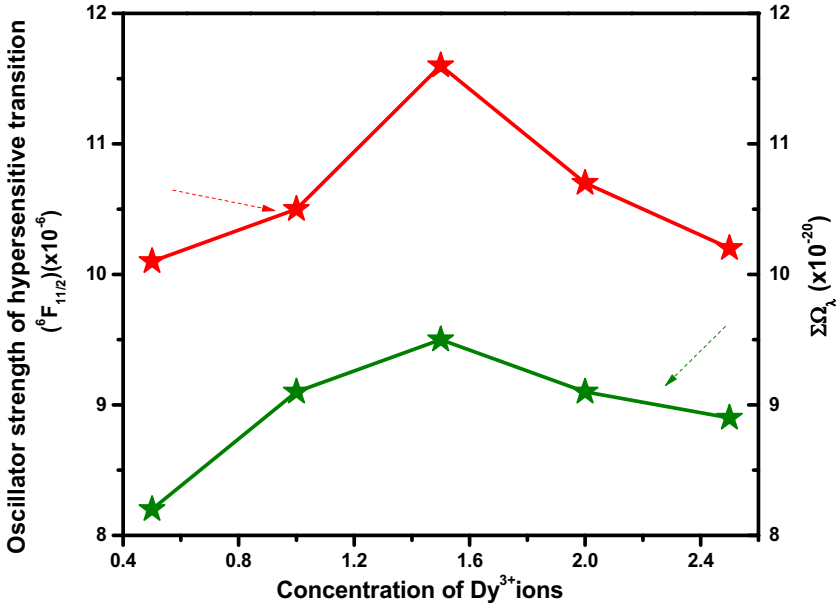


Fig. 2. Variation of intensity of hypersensitive transition and $\Sigma\Omega_\lambda$ with Dy^{3+} ions concentration in ZnPbWTe glass.

wavelength to record the PL spectra. Fig. 4 shows the PL spectra of the as prepared glasses. The spectra contains two relatively intense emission bands corresponding to the transitions $^4F_{9/2} \rightarrow {}^6H_{15/2}$ (483 nm) and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ (575 nm). The electronic dipole transition, $^4F_{9/2} \rightarrow {}^6H_{13/2}$ (575 nm) being hypersensitive in nature obeys selection rules $\Delta J = \pm 2$ and $\Delta L = \pm 2$ and is identified in yellow region of the emission spectra [35]. Ligand environment has strong influence on the intensity of the hypersensitive transition ($^4F_{9/2} \rightarrow {}^6H_{13/2}$) whereas magnetic dipole transition $^4F_{9/2} \rightarrow {}^6H_{15/2}$ (483 nm) hardly varies with the environment [36]. Considerable change in the relative intensities of emission bands (yellow and blue) is observed in PL spectra with respect to Dy^{3+} ion concentration. It can be seen from Fig. 4 that the fluorescence intensity increases with increase in concentration up to 1 mol% of Dy^{3+} ions and beyond quenching is observed which may be due to energy transfer between Dy^{3+}- Dy^{3+} ions. Fig. 5 represents the emission intensity variation of yellow ($^4F_{9/2} \rightarrow {}^6H_{13/2}$) and blue ($^4F_{9/2} \rightarrow {}^6H_{15/2}$) transitions with Dy^{3+} ion concentration having high intensity for glass B (1 mol%). Fig. 6 shows the partial energy level diagram of Dy^{3+} doped ZnPbWTe glasses with excitation, emission and non-radiative relaxation. The $^4F_{9/2}$ level is populated by non-radiative relaxation from higher energy states as shown in Fig. 6. The radiative transitions arising from $^4F_{9/2}$ level results intense blue and yellow emissions. The intense emission bands in the PL spectra (Fig. 4) were observed due to small energy gap (21,000 cm^{-1}) between $^4F_{9/2}$ level and levels above it (Fig. 6) [37].

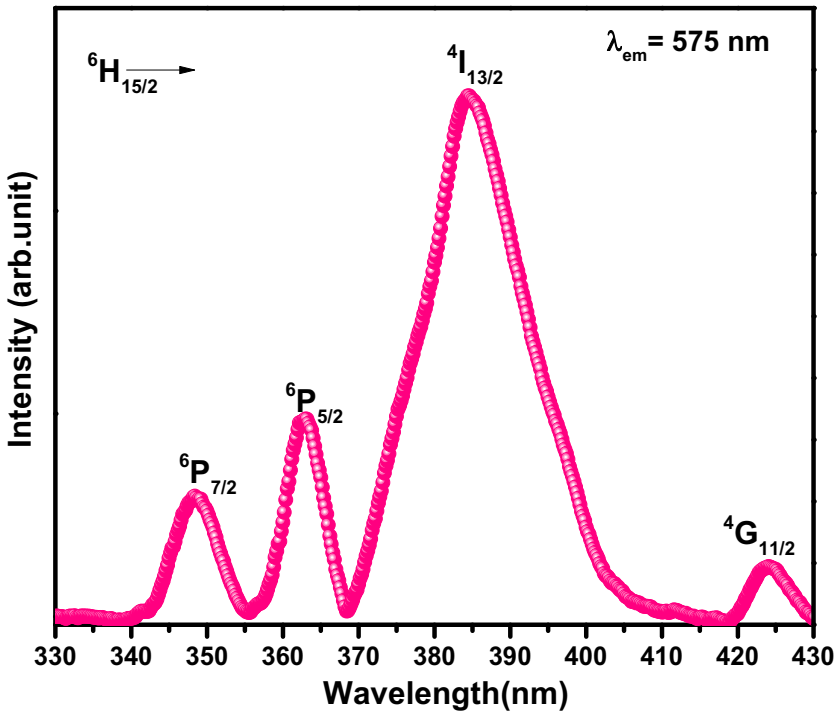


Fig. 3. Excitation spectrum of 1 mol% of Dy^{3+} ions (glass B) in ZnPbWTe glass.

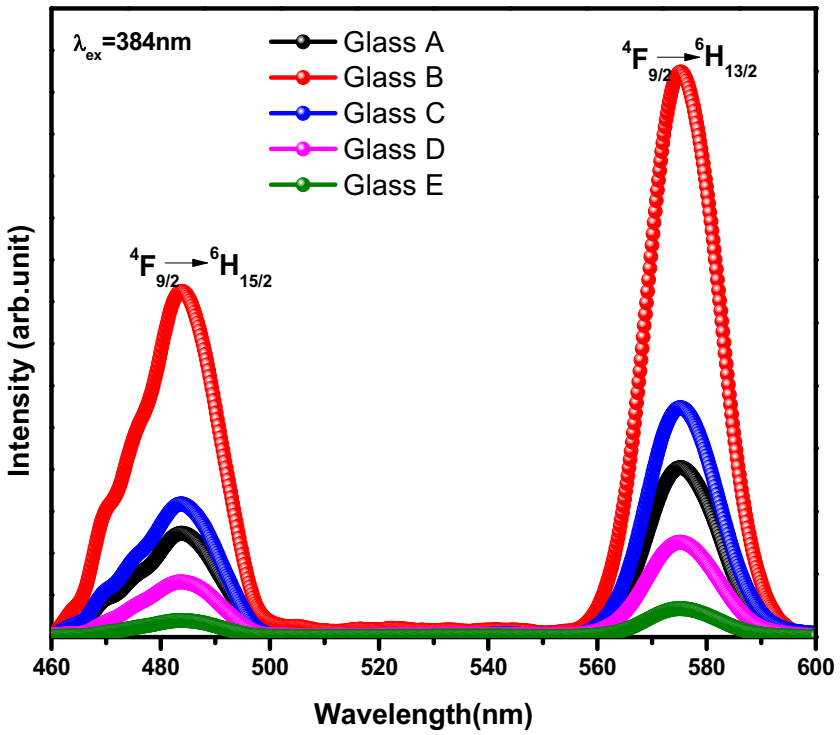


Fig. 4. Emission spectra of Dy^{3+} ions in ZnPbWTe glasses.

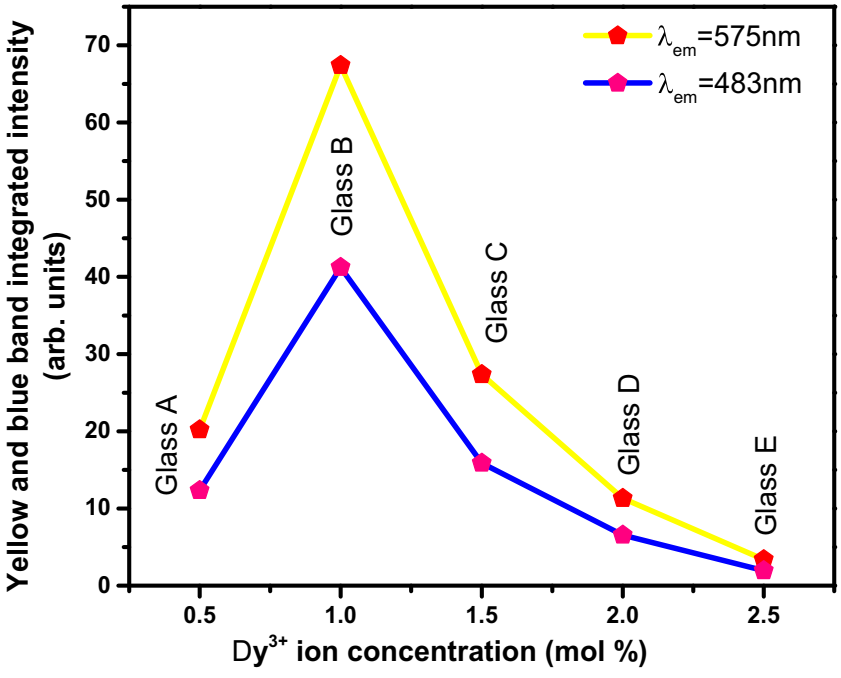


Fig. 5. Emission intensity variation of Yellow ($^4F_{9/2} \rightarrow {}^6H_{13/2}$) and blue ($^4F_{9/2} \rightarrow {}^6H_{15/2}$) transitions of Dy^{3+} ions in ZnPbWTe glasses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Radiative properties

The radiative properties such as radiative transition probability (A_R), total transition probability (A_T) and branching ratio (β_R) for the emission transitions $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ are estimated by employing J-O theory and are given in Table 3. All these radiative parameters for the titled glasses have been evaluated by using suitable formulae given in literature [38]. It is well known that for potential laser transition $\beta_R \geq 0.5$. From Table 3 it is observed that for the titled glasses, $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transitions are having higher β_R value than $^4F_{9/2} \rightarrow {}^6H_{15/2}$ transition. The experimental branching ratios (β_{exp}) are evaluated by taking the ratio of the relative intensities of individual peaks to the total intensity of emission peaks and are represented in Table 4. The non-radiative relaxation from the $^4F_{9/2}$ level of Dy^{3+} ions in the prepared ZnPbWTe glasses are responsible for the difference in the values of (β_R) and (β_{exp}). Another important parameter to find the suitable lasing action in RE ion doped glasses is stimulated emission cross-section (σ_{se}). To get good laser transition, the value of σ_{se} must be

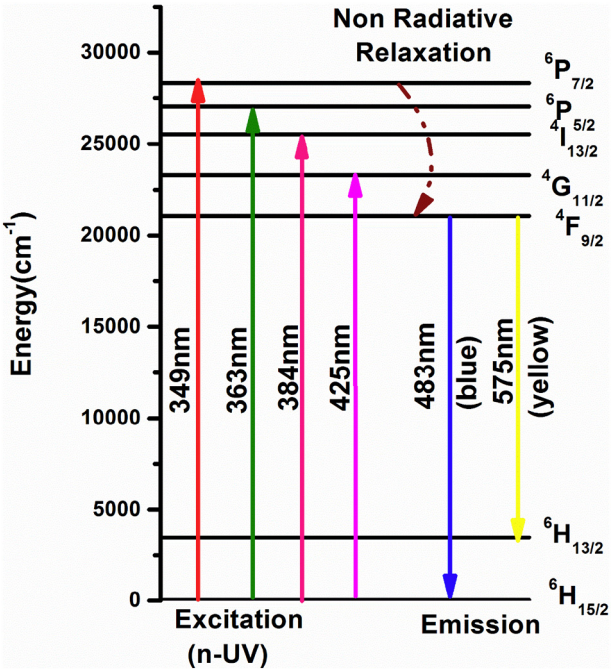


Fig. 6. Energy level diagram of 1 mol% of Dy^{3+} ions (glass B) in ZnPbWTe glass.

Table 3
Transition probabilities (A_R) (s^{-1}), luminescence branching ratios (β_R) and total transition probability (A_T) (s^{-1}) for the observed emission transitions for Dy^{3+} ions in ZnPbWTe glasses.

Transition	A_R	β_R	A_T
Glass A			
$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2425.60 ± 0.8	0.5460 ± 0.002	4210.89 ± 0.11
$^4F_{9/2} \rightarrow {}^6H_{15/2}$	1016.57 ± 0.9	0.2414 ± 0.001	
Glass B			
$^4F_{9/2} \rightarrow {}^6H_{13/2}$	3048.76 ± 0.3	0.5682 ± 0.004	5461.75 ± 0.25
$^4F_{9/2} \rightarrow {}^6H_{15/2}$	1505.78 ± 0.7	0.2757 ± 0.001	
Glass C			
$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2851.59 ± 0.4	0.5643 ± 0.009	5053.26 ± 0.13
$^4F_{9/2} \rightarrow {}^6H_{15/2}$	1302.44 ± 0.8	0.2577 ± 0.004	
Glass D			
$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2733.91 ± 0.6	0.5611 ± 0.002	4872.72 ± 0.37
$^4F_{9/2} \rightarrow {}^6H_{15/2}$	1263.24 ± 0.4	0.2592 ± 0.001	
Glass E			
$^4F_{9/2} \rightarrow {}^6H_{13/2}$	2601.33 ± 0.2	0.5519 ± 0.007	4548.28 ± 0.41
$^4F_{9/2} \rightarrow {}^6H_{15/2}$	1099.65 ± 0.5	0.2418 ± 0.006	

large. The σ_{se} parameter obtained for $^4F_{9/2} \rightarrow {}^6H_{15/2}$ and $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transitions in the titled glasses are given in Table 4 which are in good agreement with the values reported for other hosts [39,40]. For the as prepared glasses, gain band width ($\sigma_{se} \times \Delta\lambda_p$) and optical gain ($\sigma_{se} \times \tau_R$) have also been evaluated and presented in Table 4. From Table 4 it is observed that for glass B (1 mol% Dy^{3+}) the values of β_R and σ_{se} are high, indicating the utility of these glasses as promising material for laser action. In addition to this higher $\sigma_{se} \times \Delta\lambda_p$ and $\sigma_{se} \times \tau_R$values make glass B as a potential candidate for making optical fiber and optical amplifiers.

3.4. Fluorescence decay spectral analysis

Fig. 7 represents the decay profiles of Dy^{3+} doped ZnPbWTe glasses under 384 nm excitation and 575 nm emission. From Fig. 7 it is observed that decay profiles are exponential for lower concentration and non-exponential for higher concentration. This can be attributed to the energy transfer between Dy^{3+}-Dy^{3+} ions. The experimental lifetime (τ_{exp}) of $^4F_{9/2}$ fluorescent level has been estimated by taking first e-fold time of the decay curves. The relation between experimental lifetime τ_{exp}, radiative lifetime τ_R and non-radiative relaxation rate W_{NR} is given by the following expression

$$\frac{1}{\tau_{exp}} = \frac{1}{\tau_R} + W_{NR}$$

here, τ_{exp} is obtained from decay curves whereas τ_R can be estimated from *J*-O theory. τ_{exp} and τ_R values for the synthesized glasses are presented in Table 5. From Table 5 it is noticed that τ_{exp} values for the titled glasses are lower than τ_R values which may be due to non-radiative relaxation of Dy^{3+} ions as well as interaction between Dy^{3+} ions and OH−- stretching. In order to predict the quality of laser host material, quantum efficiency of the as prepared glasses are evaluated by using following expression

$$\eta = \frac{\tau_{exp}}{\tau_R} \times 100$$

For all the titled glasses the value of η has been evaluated and is tabulated in Table 5. Among all the ZnPbWTe glasses it is observed that glass B exhibits highest quantum efficiency around 87%.

3.5. Inokuti-Hirayama model analysis

Inokuti-Hirayama (I-H) model is an interesting model and applicable to evaluate the dominant mechanism of interaction for energy transfer between the $4f^n$ energy levels of active ions. By fitting the experimental decay data to I-H model, the non-exponential behaviour of the decay curves can be analyzed [41]. The non-exponential nature of decay curves arises at higher concentration of RE ions because at higher concentration RE ions come close to each other leading to remarkable interaction and hence energy transfer takes place from excited RE ion to unexcited RE ion.

$$I = I_0 \exp\left\{-\frac{t}{\tau_0} - Q\left(\frac{t}{\tau_0}\right)^{\frac{3}{S}}\right\}$$

here t is the time after excitation, τ_0 is the intrinsic decay time of the isolated donors, where no energy transfer takes place and the decay curves exhibit perfect exponential nature. Q is the energy transfer parameter which is given by the following expression

$$Q = \frac{4\pi}{3}\Gamma\left(1 - \frac{3}{S}\right)N_0 R_0{}^3$$

here, N_0 is concentration of acceptor ions which is almost equal to the total concentration of RE ions. Critical energy transfer distance R_0 can be obtained as the separation between donor and acceptor when the rate of energy transfer between donor and acceptor is equal to rate of intrinsic decay τ_0^{-1}.

The magnitude of energy transfer parameter Q depends on S as well as gamma function Γ. The values S = 6, 8 and 10 correspond to dipole-dipole (Γ = 1.77), dipole-quadrupole (Γ = 1.43) and quadrupole-quadrupole (Γ = 1.30) interaction respectively. For the as-prepared glasses under investigation, decay curves for higher concentration are well fitted with S = 6 as shown in Fig. 8. Due to increase in the concentration of RE ions, R_0 decreases hence Q parameter increases which results in faster fluorescence decay and the values (Q and R_0) for higher concentration (glass C, D, E) are depicted in Table 5.

3.6. CIE chromaticity coordinates

The generation of white light by Dy^{3+} ions in ZnPbWTe glasses has been analyzed in the framework of the CIE chromaticity color coordinates. The color coordinates for all the as prepared glasses were

Table 4
Emission peak wavelength (λ_P) (nm), effective band widths ($\Delta\lambda_P$) (nm), measured and experimental branching ratios (β_R & β_{exp}), stimulated emission cross-sections (σ_{se}) (cm^2), gain band width ($\sigma_{se} \times \Delta\lambda_P$) ($cm^3$) and optical gain ($\sigma_{se} \times \tau_R$) ($cm^2s$) for emission transitions for Dy^{3+} ions in ZnPbWTe glasses.

Spectral parameters	Glass A	Glass B	Glass C	Glass D	Glass E
$^4F_{9/2} \rightarrow {}^6H_{15/2}$ (blue) λ_P = 483 nm					
$\Delta\lambda_P$	18.1	18.39	18.57	16.94	13.21
β_R	0.241 ± 0.001	0.275 ± 0.001	0.257 ± 0.004	0.259 ± 0.001	0.241 ± 0.006
β_{exp}	0.405	0.495	0.395	0.375	0.374
σ_{se} ($\times 10^{-22}$)	7.21 ± 0.9	11.4 ± 0.7	7.34 ± 0.8	7.52 ± 0.4	8.26 ± 0.5
$\sigma_{se} \times \Delta\lambda_P$ ($\times 10^{-28}$)	13.1	18.7	13.5	12.8	10.9
$\sigma_{se} \times \tau_R$ ($\times 10^{-25}$)	1.71	2.09	1.42	1.66	1.69
$^4F_{9/2} \rightarrow {}^6H_{13/2}$ (yellow) λ_P = 575 nm					
$\Delta\lambda_P$	14.7	15.18	15.49	15.19	15.19
β_R	0.546 ± 0.002	0.568 ± 0.004	0.564 ± 0.009	0.561 ± 0.002	0.551 ± 0.007
β_{exp}	0.571	0.577	0.567	0.555	0.536
σ_{se} ($\times 10^{-22}$)	36.6 ± 0.8	49.6 ± 0.3	37.7 ± 0.4	36.0 ± 0.6	33.5 ± 0.2
$\sigma_{se} \times \Delta\lambda_P$ ($\times 10^{-28}$)	61.1	74.9	58.0	54.8	51.0
$\sigma_{se} \times \tau_R$ ($\times 10^{-25}$)	8.66	9.07	7.42	7.89	6.87

calculated using emission spectra based on the Commission International deI'Eclairage (CIE) 1931 standards [42]. The CIE co-ordinates are considerably stable against the enhanced intensity of emission bands by increasing the Dy^{3+} ion concentration for the titled glasses and are represented in Table 6. The glass B represents excellent chromaticity coordinates (0.312, 0.324) in white light region under 384 nm excitation wavelength. The CIE coordinates of glass B are shown in Fig. 9. It is evident that these coordinates are very close to the standard white light illumination point (0.333, 0.333) which possess white light emission in the chromaticity diagram. Further, white light chromaticity coordinates exhibited by optimized glass B are very close to the chromaticity coordinates of commercial pc-LED (0.32, 0.32) and NTSC white light (0.310, 0.316) [43]. From these results it is observed that the present glasses are aptly suitable for white light generation. The color correlated temperature (CCT) is another important parameter to investigate the quality of light source. CCT is defined as the temperature of planckian black-body radiator whose color is nearest to the color of white light source [41]. The following equation given by Mccamy has been used to calculate the value of CCT [44].

Table 5
Experimental lifetime (τ_{exp}) (μs), energy transfer parameter (Q),critical transfer distance (R_0) (Å), radiative lifetime (τ_R), quantum efficiencies (η) and non-radiative decay rates (W_{NR}) for $^4F_{9/2} \rightarrow {}^6H_{15/2}$ transition of Dy^{3+} in ZnPbWTe glasses.

Name of sample	τ_{exp}	Q	R_0	τ_R	η (%)	W_{NR}
Glass A	185	–	–	237	78	1200
Glass B	160	–	–	183	87	790
Glass C	146	2.9	3.10	197	74	1770
Glass D	134	3.2	2.69	219	61	2900
Glass E	113	4.5	2.66	205	55	3970

$$CCT = -449n^3 + 3525n^2 - 6823.2n + 5520.3$$

where $n = \frac{X - X_E}{Y - Y_E}$, $X_E = 0.332$, $Y_E = 0.186$ is the epicenter. For all the as prepared glasses, the CCT values (for λ_{ex} = 384 nm) have been calculated and are given in Table 6. From Table 6 it is observed that the CCT values for all the glasses under investigation are falling in the range 6013–7548 K which corresponds to cool white light region. Higher CCT values result in better visual acuity and vividness [45]. Among the as

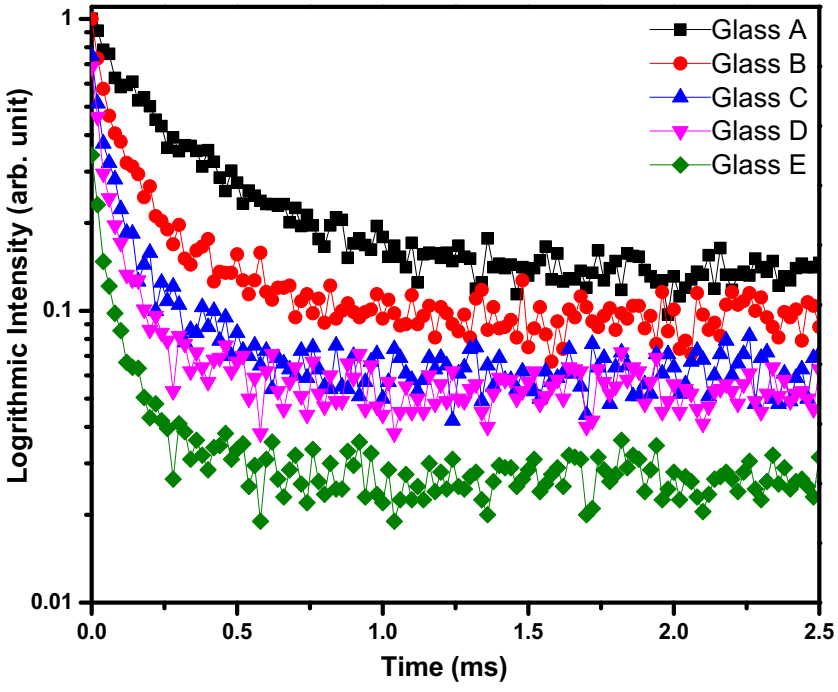


Fig. 7. Decay profiles of Dy^{3+} ions in ZnPbWTe glasses for 575 nm emission under 384 nm excitation.

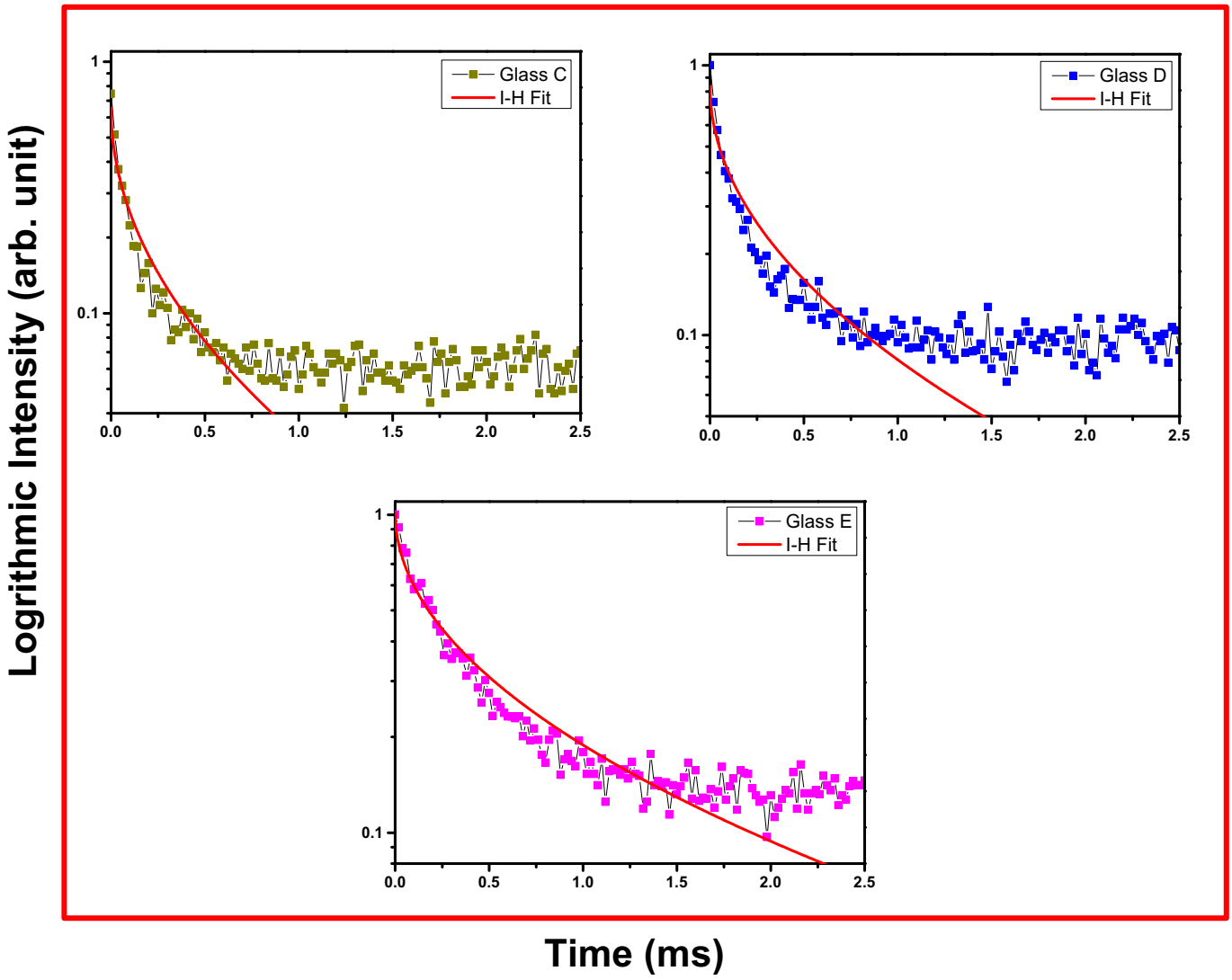


Fig. 8. I-H fitted curves (Plotted by using Origin 9.2 software) of Dy^{3+} ions in ZnPbWTe glasses for 575 nm emission under 384 nm excitation.

Table 6
Colorimetric parameters of Dy^{3+} ions in ZnPbWTe glasses underexcitation wavelength (384 nm).

Excitation wavelength	Parameters	Glass samples				
		Glass A	Glass B	Glass C	Glass D	Glass E
384 nm	(x, y)	(0.289, 0.313)	(0.312, 0.324)	(0.301, 0.316)	(0.302, 0.315)	(0.303, 0.316)
	Y/B ratio	1.63	1.63	1.72	1.72	1.74
	CCT (K)	7532	7548	7543	6112	6013

prepared glass, glass B is having highest CCT value (7548 K). The color purity can be evaluated by employing the following formula

$$\text{Color purity} = \frac{\sqrt{(X - X_{EE})^2 + (Y + Y_{EE})^2}}{\sqrt{(X_D - X_{EE})^2 + (Y_D + Y_{EE})^2}}$$

here (X, Y), (X_{EE}, Y_{EE}) and (X_D, Y_D) are the chromaticity coordinates of the color source under investigation, equal energy reference illuminant and dominant wavelength point respectively. The dominant wavelength points for glass B are shown in Fig. 9. Color purity was calculated for color coordinates (0.312, 0.324) of glass B, under 384 nm excitation wavelength. For efficient white light emission, color purity must have low value and it is observed from the present study that color purity is approximately 8×10^{-2} which makes glass B as a potential candidate for w-LEDs.

4. Conclusions

Zinc Lead Tungsten Tellurite glasses having good optical quality and doped with different concentrations of Dy^{3+} ions were prepared by conventional melt quenching technique. The luminescence properties of Dy^{3+} ions in the as prepared glasses were investigated by using absorption, emission and decay spectral measurements. The J-O theory applied to the measured oscillator strengths allowed to estimate radiative properties for the prominent fluorescent levels of the Dy^{3+} in the as prepared glasses. The room temperature luminescence spectra recorded for the glasses under investigation show two strong emissions at yellow ($^4F_{9/2} \rightarrow {}^6H_{13/2}$) and blue ($^4F_{9/2} \rightarrow {}^6H_{15/2}$) regions. Relatively higher values obtained for experimental branching ratios and stimulated emission cross-sections for $^4F_{9/2} \rightarrow {}^6H_{13/2}$ transition suggest that the as prepared glasses are potential candidates for visible laser materials. Among all the ZnPbWTe glasses, glass B (1 mol% of Dy^{3+} ions) exhibits fairly good radiative properties as well as quantum efficiency and acts as a promising luminescence material for developing visible luminescent devices. The CIE color chromaticity coordinates (x, y), CCT, color purity and Y/B ratios were also evaluated from the luminescence spectra for all the glasses. From the aforementioned results, it

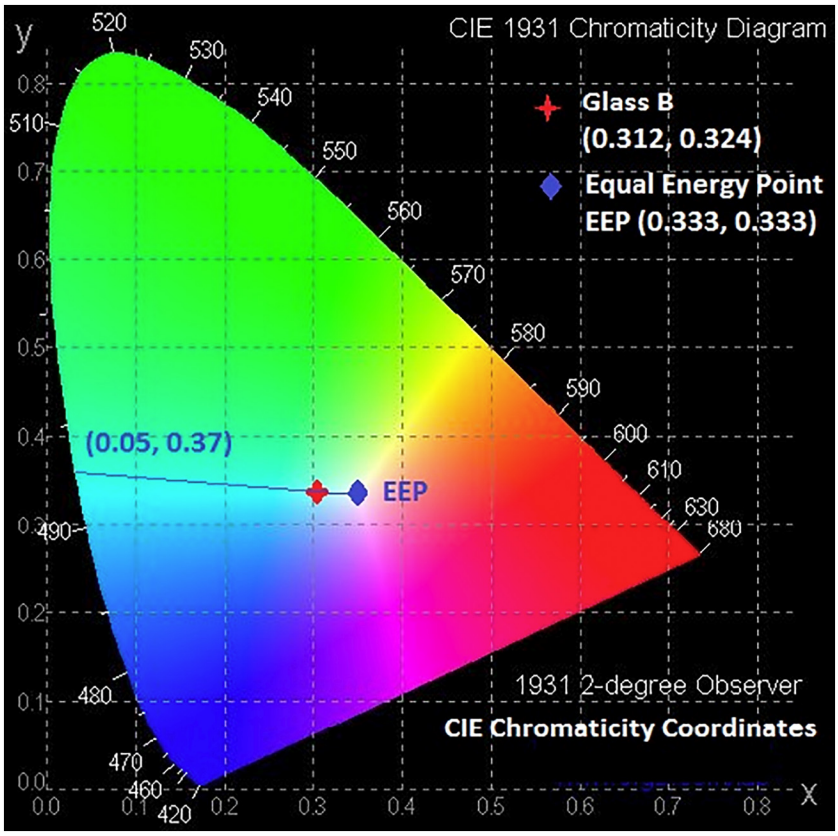


Fig. 9. CIE chromaticity coordinates of 1 mol% Dy^{3+} ions (glass B) in ZnPbWTe glass.

is concluded that among the as prepared glasses, glass B is a potential candidate for w-LEDs and lasers.

Acknowledgement

One of the authors, Prof. A. S. Rao is thankful to Department of Science and Technology (DST), Govt. of India, New Delhi for the award of a major research project (File Number: EMR/2016/007766) to him under DST-SERB.

References

[1] K. Jha, M. Jayasimhadri, Structural and emission properties of Eu^{3+}-doped alkaline earth zinc-phosphate glasses for white LED applications, J. Am. Ceram. Soc. (2017) 1–10, http://dx.doi.org/10.1111/jace.14668.
[2] A.S. Rao, Y.N. Ahammed, R.R. Reddy, T.V.R. Rao, Spectroscopic studies of Nd^{3+} doped alkali fluoroborophosphate glasses, Opt. Mater. 10 (1998) 245–252.
[3] N. Deopa, A.S. Rao, Photoluminescence and energy transfer studies of Dy^{3+} ions doped lithium lead alumino borate glasses for w-LED and laser applications, J. Lumin. 192 (2017) 832–841.
[4] S.A. Saleem, B.C. Jamalaiah, M. Jayasimhadri, A.S. Rao, K. Jang, L.R. Moorthy, Luminescent studies of Dy^{3+} ion in alkali lead telluro fluoroborate glasses, J. Quant. Spectrosc. Radiat. Transf. 112 (2011) 78–84.
[5] N. Deopa, A.S. Rao, Spectroscopic studies of Sm^{3+} ions activated lithium lead alumin borate glasses for visible luminescent device applications, Opt. Mater. 72 (2017) 31–39.
[6] S. liu, G. Zhao, X. Lin, H. Ying, J. Liu, J. Wang, G. Han, White luminescence of Tm –Dy ions co-doped alumina boro silicate glasses under UV light excitation, J. Solid State Chem. 181 (2008) 2725–2730.
[7] C.B. Annapurna Devi, S. Mahamuda, M. Venkateswarlu, K. Swapna, A.S. Rao, G. Vijaya Prakash, Dy^{3+} ions doped single and mixed alkali fluoro tungsten tellurite glasses for LASER and white LED applications, Opt. Mater. 62 (2016) 569–577.
[8] K. Swapna, S.K. Mahamuda, A.S. Rao, M. Jayasimhadri, T. Sasikala, L.R. Moorthy, Optical absorption and luminescence characteristics of Dy^{3+} doped Zinc Alumino Bismuth Borate glasses for lasing materials and white LEDs, J. Lumin. 139 (2013) 119–124.
[9] N. Deopa, A.S. Rao, Photoluminescence and energy transfer studies of Dy^{3+} ions doped lithium lead alumino borate glasses for w-LED and laser applications, J. Lumin. 192 (2017) 832–841.
[10] A.M. Babu, B.C. Jamalaiah, J. Suresh Kumar, T. Sasikala, L.R. Moorthy, Spectroscopic and photoluminescence properties of Dy^{3+}-doped lead tungsten tellurite glasses for laser materials, J. Alloys Compd. 509 (2011) 457–462.
[11] G. Poirier, F.S. Ottoboni, F.C. Cassanjes, A. Remonte, Y. Messaddeq, S.J.L. Ribeiro, Redox behavior of molybdenum and tungsten in phosphate glasses, J. Phys. Chem. B 112 (2008) 4481–4487.
[12] M. Ohtsuki, R. Tamura, S. Takeuchi, S. Yoda, T. Ohmura, Hard metallic glass of tungsten-based alloy, Appl. Phys. Lett. 84 (2004) 4911–4913.
[13] M.A. Engelhardt, S.S. Jaswal, D.J. Sellmyer, Photoemission and electronic structure of tungsten-based metallic glasses and alloys, Phys. Rev. B 44 (1991) 12671–12679.
[14] S.M. Kaczmarek, $Li_2B_4O_7$ glasses doped with Cr, Co, Eu and Dy, Opt. Mater. 19 (2002) 189–194.
[15] K. Jha, M. Jayasimhadri, Spectroscopic investigation on thermally stable Dy^{3+} doped zinc phosphate glasses for white light emitting diodes, J. Alloys Compd. 688 ((2016) 833–840.
[16] G.A. Kumar, A. Martinez, E. Mejia, C.G. Eden, Fluorescence and up conversion spectral studies of Ho^{3+} in alkali bismuth gallate glasses, J. Alloys Compd. 365 (2004) 117–120.
[17] N. Deopa, A.S. Rao, Spectroscopic studies of single near ultraviolet pumped Tb^{3+} doped lithium lead alumino borate glasses for green lasers and tricolor w-LEDs, J. Lumin. 194 (2018) 56–63.
[18] C.M. Reddy, G.R. Dillip, B.D.P. Raju, Spectroscopic and photoluminescence characteristics of Dy^{3+} ions in lead containing sodium fluoroborate glasses for laser materials, J. Phys. Chem. Solids 72 (2011) 1436–1441.
[19] S. Liu, G. Zhao, H. Ying, J. Wang, G. Han, Eu/Dy ions co-doped white light luminescence zinc–alumino boro silicate glasses for white LED, Opt. Mater. 31 (2008) 47–50.
[20] M. Jayasimhadri, K. Jang, H.S. Lee, B. Chen, S.S. Yi, J.H. Jeong, White light generation from Dy^{3+} doped ZnO-B_2O_3-P_2O_5 glasses, J. Appl. Phys. 106 (2009) 013105.
[21] C.R. Kesavulu, C.K. Jayasankar, White light emission in Dy^{3+}-doped lead fluorophosphate glasses, Mater. Chem. Phys. 130 (2011) 1078–1085.
[22] J.L. Cai, R. Li, C. Zhao, S. Tie, X. Wan, J. Shen, White light emission and energy transfer in Dy^{3+}/Eu^{3+} co-doped aluminoborate glass, Opt. Mater. 34 (7) (2012) 1112–1115.
[23] K. Swapna, Sk Mahamuda, A. Srinivasa Rao, M. Jayasimhadri, T. Sasikala, L.R. Moorthy, Visible fluorescence characteristics of Dy^{3+} doped Zinc alumino bismuth borate glasses for optoelectronic devices, Ceram. Int. 39 (2013) 8459–8465.
[24] W.T. Carnall, P.R. Fields, K. Rajnak, Electronic energy levels in the trivalent lanthanide aquo Ions. I. Pr^{3+}, Nd^{3+}, Pm^{3+}, Sm^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, and Tm^{3+}, J. Chem. Phys. 49 (1968) 4424–4442.
[25] P. Babu, C.K. Jayasankar, Spectroscopic properties of Dy^{3+} ions in lithium borate

and lithium fluoroborate glasses, Opt. Mater. 15 (2000) 65–79.
[26] B.R. Judd, Optical absorption intensities of rare-earth ions, Phys. Rev. B 127 (1962) 750–761.
[27] G.S. Ofelt, Intensities of crystal spectra of rare-earth ions, J. Chem. Phys. 37 (1962) 511–520.
[28] G. Dominiak-Dzik, W. Ryba-Romanowski, M.N. Palatnikov, N.V. Sidorov, V.T. Kalnnikov, Dysprosium-doped $LiNbO_3$ crystal, optical properties and effects of temperature on fluorescence dynamics, J. Mol. Struct. 704 (2004) 139–144.
[29] W.F. Krupke, Induced-emission cross sections in neodymium laser glasses, IEEE J. Quantum Electron. 10 (1974) 450–457.
[30] W. Ryba- Romanowski, G.D. Dzik, P. Solarz, R. Lisiecki, Transition intensities and excited state relaxation dynamics of Dy^{3+} ions in crystals and glasses, Opt. Mater. 31 (2009) 1547–1554.
[31] Y.B. Shin, J. Heo, Mid-infrared emissions and multiphonon relaxation in Dy3+ doped chalcohalide glasses, J. Non-Cryst. Solids 253 (1999) 23–29.
[32] Y.C. Ratnakaram, D.T. Naidu, A. Vijayakumar, J.L. Rao, Studies on optical absorption and luminescence properties of Dy3+ doped mixed alkali borate glasses, Opt. Mater. 27 (2004) 409–417.
[33] J. Pisarska, Optical properties of lead borate glasses containing Dy^{3+} ions, J. Phys. Condens. Matter 21 (2009) 285101.
[34] C.K. Jorgensen, R. Reisfeld, Judd-Ofelt parameters and chemical bonding, J. Less-Common Met. 93 (1983) 107–112.
[35] S. Tanabe, T. Ohyagi, N. Soga, T. Hanada, Compositional dependence of Judd-Ofelt parameters of Er^{3+} ions in alkali–metal borate glasses, Phys. Rev. B 46 (1992) 3305–3310.
[36] G.V.L. Reddy, L.R. Moorthy, B.C. Jamalaiah, T. Sasikala, Preparation, structural and luminescent properties of $YAl_3(BO_3)_4$:Dy^{3+}phosphor for white light-emission under UV excitation, Ceram. Int. 39 (2013) 2675–2682.
[37] Z. Yang, W. Chen, L. Luo, Dy^{3+}-doped Ge-Ga-Sb-Se glasses for 1.3 μm optical fiber amplifiers, J. Non-Cryst. Solids 351 (2005) 2513–2518.
[38] A.S. Rao, B. Rupa, V. Rao, M.V.V.K. Prasad, J.V.S. Kumar, M. Jayasimhadri, J.L. Rao, R.P.S. Chakradhar, Spectroscopic and optical properties of Nd^{3+} doped fluorine containing alkali and alkaline earth zinc-alumino phosphate optical glasses, Physica B 404 (2009) 3717–3721.
[39] M. Jayasimhadri, L.R. Moorthy, K. Kojima, K. Yamamoto, Wada Noriko, Wada Noriyuki, Optical properties of Dy^{3+} ions in alkali telluroflurophosphate glasses for laser material, J. Phys. D 39 (2006) 635–641.
[40] B.C. Jamalaiah, J.S. Kumar, T. Suhasini, K.W. Jang, H.S. Lee, H.J. Choi, L.R. Moorthy, Visible luminescence characteristics of Dy^{3+} J doped LBTAF glasses, J. Alloys Compd. 474 (2009) 382–387.
[41] M. Inokuti, F. Hirayama, Influence of energy transfer by the exchange mechanism on donor luminescence, J. Chem. Phys. 43 (1965) 1978–1988.
[42] E.F. Schubert, Light Emitting Diodes, 2nd edition, Cambridge University Press, 2006 292 pp. (chapter 17).
[43] T.W. Kuo, W.R. Liu, T.M. Chen, High color rendering white light-emitting-diode illuminator using the red-emitting Eu^{2+} activated CaZnOS phosphors excited by blue LED, Opt. Express 18 (2010) 8187–8192.
[44] C.S. McCamy, Correlated color temperature as an explicit function of chromaticity coordinates, Color. Res. Appl. 17 (1992) 142–144.
[45] S.M. Berman, M. Navvab, M.J. Martin, J. Sheedy, W. Tithof, Light. Res. Technol. 38 (2006) 41–52.

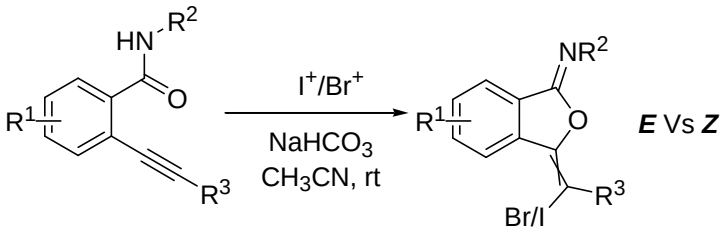
International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Product Geometry Analysis of Cyclic Imidates using Molecular Mechanics

Dhirendra Brahmchari and Saurabh Mehta*
Department of Applied Chemistry,
Delhi Technological University, Bawana Road, Delhi, India

ABSTRACT

Under mildly basic conditions the o-(1-alkynyl)benzamides undergo electrophilic cyclization yielding cyclic imidates. The products formed were found to possess exclusive E-geometry across the exocyclic C=C double bond which was confirmed by single crystal X-ray crystallographic data and NMR spectroscopic data. The observed stereoselectivity of the reaction was investigated, and the thermodynamic stability of the cyclic imidates having exocyclic C=C bond was calculated using molecular mechanics (MM2) methods. Various parameters like stretch, bend, torsion, Van der Waals forces and dipole were evaluated to calculate the total energy. The MM2 results obtained were in close agreement with the experimental findings, thereby, confirming that the cyclic imidates having E-geometry across the C=C bond are thermodynamically more stable than their Z-geometric counterparts.



R^1 = H, OMe; R^2 = Me, Ph, Bn; R^3 = H, TMS, Ph, etc.
E = I, Br

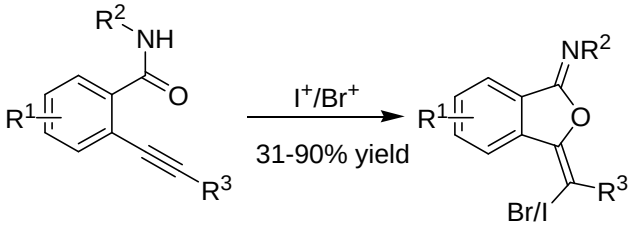
KEYWORDS: Calculations, Cyclic Imidate, MM2, Molecular Mechanics

INTRODUCTION

Heterocyclic frameworks having diverse substitution patterns are found in several naturally occurring as well as medicinally important compounds, and have displayed great importance in diverse areas, including medicinal chemistry, materials science, *etc*.[1] Amongst the vast variety of heterocyclic compounds known, cyclic imidates have gained attention of the chemists owing to their useful biological activities.[2–7] Several research groups have developed various interesting approaches for the synthesis of cyclic imidates which includes metal mediated cyclizations,[8–10] base mediated cyalizations[11–13] *etc*.[14]

We and others have reported that under the mild reaction conditions 2-(1-alkynyl)-benzamides undergoes regio- and stereoselective cyclization affording corresponding cyclic imidates in high yields and the cyclization occurred regioselectively *via* the *O*-atom of the amide group (**Figure 1**).[11]

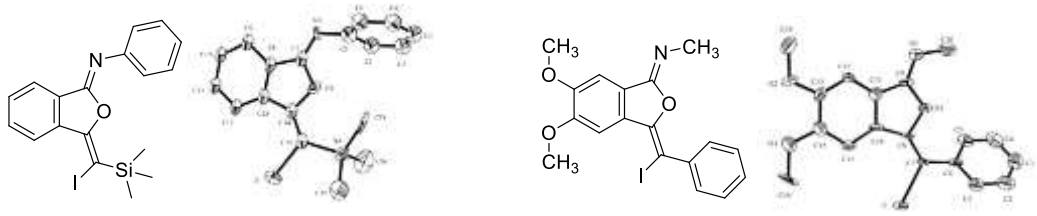
Figure 1. Regio- and Stereoselective cyclization of 2-(1-alkynyl)-Benzamides to form Cyclic Imidates



International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

The alkynyl amides were found to cyclize *via* 5-*exo*-*dig* cyclization in the presence of suitable electrophile and the resulting cyclic imidates formed were found to possess the exclusive *E*-geometry across the exocyclic C=C bond (confirmed using single-crystal X-ray crystallographic data and NMR spectroscopic data, **Figure 2**).

Figure 2. ORTEP diagrams of cyclic imidates formed confirming the formation of *E*-geometric isomer.[11]

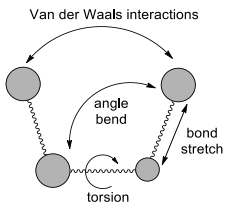


A similar finding (*O*-cyclization of 2-(1-alkynyl)-benzamides) was reported by Schlemmer *et. al.*[13] Moreover, all cyclic imidates formed during this methodology were found to exhibit *E*-geometry across the exocyclic C=C bonds. Since the methodology could have resulted in the formation of both *E* and *Z*-isomers, so we became interested to find the factors responsible for the selective formation of the *E*-product. To justify the formation of the actual products during the experimentation, we did the additional computational experiments on the cyclic imidates formed.

Over the years the developments in computational chemistry have helped researchers in exploring new avenues and crossing the traditional boundaries that separated chemistry from other research areas of biology, physics, *etc*. In Molecular modelling computer simulations are used for describing the behaviour of molecules and it presents a reasonable alternative in situations where the actual laboratory exploration may not be practical. Molecular mechanics is a classical physics approach that is routinely used by chemists.[15,16] It comprises of a set of potential energy functions for modelling the behaviour of matter. Molecular modelling methods have been valuable aids to the synthetic chemists.[17] Various parameters and potential energy functions used for the evaluation of molecular mechanics are collectively known as a "force-field" and are governed by the fundamental principles of molecular mechanics.[18–20] Popular force fields include MM2, MM3, and MMFF.[21,22] The MM2 force field was developed by N. L. Allinger[23] and it was found to be suitable for various organic compounds. In molecular mechanics, a force field (bond angles, lengths, electrostatic properties, torsional parameters, Van der Waals interactions, *etc*.) is used for analyzing the steric energy of the molecule and amending the conformation to minimize the steric energy. MM2 calculations are very fast and can be used for calculation of molecular dynamics for even bulky molecules such as enzymes. The parameters of MM2 were postulated by N. L. Allinger[23] and others.[18,24,25] In molecular mechanics, atoms and bonds are considered as per the spheres and spring model where the ability to stretch, bend, and the twist is defined using the Hooke's Law. The non-bonded interactions include the electrostatic attraction, steric repulsion, van der Waals attraction,[26] *etc*. The significant energy contribution is from the torsional energy and from the 1,4-VDW interactions. The summation of various interactions such as Stretch Energy (E_{str}), Bending Energy (E_{bend}), Dipole-Dipole Energy (E_{el}), Van der Waals Energy (E_{vdw}), Stretch-Bend Energy ($E_{str\text{-}bend}$), Torsion Energy (E_{tors}), describes the total energy (E_{total}) of the molecule:

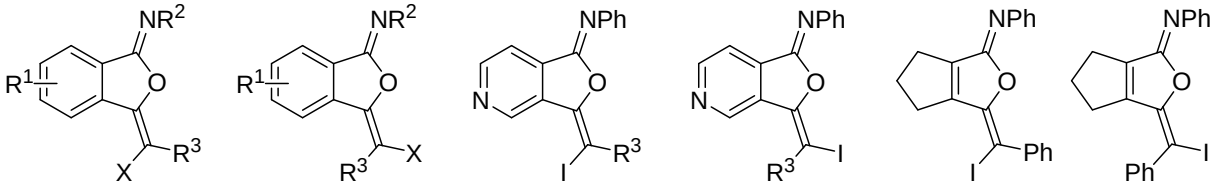
$$E_{total} = E_{str} + E_{bend} + E_{str\text{-}bend} + E_{vdw} + E_{tors} + E_{el}$$

Figure 3. Major interactions in a molecule



International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Figure 4. Various cyclic imidates (1-13) screened using the MM2 molecular modelling



1-9 **1-9** **10-12** **10-12** **13** **13**

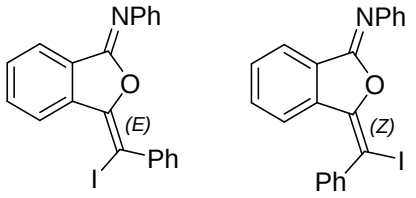
R^1 = H, OMe,
R^2 = Me, Ph, Bn
R^3 = Ph, TMS, *n*-C_8H_{17}, cyclohexenyl,
X = Br, I

R^3 = Ph, *n*-C_5H_{11}, cyclohexenyl

RESULTS AND DISCUSSION

Application of the Molecular mechanics for the calculation of the thermodynamic stability of various geometric cyclic imidates that could be formed during the electrophilic cyclization of *o*-(1-Alkynyl)benzamides provides with the interesting insights. Using the experimental procedure for the calculation of MM2 results for the cyclic imidate **1** (both *E* & *Z*-geometric isomers), the results and the parameters obtained were summarized as given below (**Table 1**). It was observed that the MM2 results were in close agreement with the experimental outcomes. MM2 results confirmed that the product having the *E*-geometry across the exocyclic C=C product to be more stable than the product having the *Z*-geometry having an energy difference {**1**(*Z*)-**1**(*E*)} of 1.4077 kcal/mol (calculated, **Table 1**).

Table 1: Output tab displaying energy calculations of various parameters from an MM2 job for imidate 1.



Energy Term	**1**(*E*)-isomer	**1**(*Z*)-isomer
Stretch	2.3617	2.1053
Bend	21.1967	20.0511
Stretch-Bend	0.2384	0.1250
Torsion	-16.9884	-12.7264
Non-1,4 VDW	5.3041	2.6720
1,4 VDW	22.5889	22.9525
Dipole/Dipole	0.4099	1.3394
Total energy	35.1113 kcal/mol	36.5190 kcal/mol

Intrigued by these results, we performed the MM2 calculations for both the geometrical isomers (*E* & *Z*) of all the cyclic imidates (**1-13**) and the results are summarized in **Table 2**. The MM2 results obtained for all the cyclic imidates showed that in all the cases, the total energy values for the *E*-geometric product were less as compared to the *Z*-geometric product, which is in agreement with the experimental results. Therefore, during the electrophilic cyclization of 2-(1-alkynyl)benzamides the *E*-geometric product was the chemically favoured and thermodynamically stable product. The results of the comparison of various MM2 calculation parameters

for both the geometric products were summarized in **Table 2**. After obtaining the MM2 calculation results, the energy difference between the thermodynamic stability of cyclic imidates (**1**-**13**) and their corresponding isomeric counterparts was found to be greatest in case of imidate **11** (9.9538 kcal/mol) while it was found to be least in case of imidate **6** (0.0003 kcal/mol). The details of the energy difference of the corresponding geometrical isomers of various cyclic imidates were summarized in **Table 2**. A comparison of the total energy of both (*E* and *Z*) geometric isomers is shown in **Figure 5,** and the calculated energy difference between corresponding geometrical isomers (*E*-*Z*) is displayed in **Figure 6**.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Table 2: Comparison of the MM2 energy parameters of *E*- & *Z* cyclic imidates (1-13)

Cyclic imidate	Stretch energy		Bend energy		Torsion		Non-1,4 VDW		1,4 VDW		Dipole/ Dipole		Total energy		Energy difference (kcal/mol)
	E	*Z*	*E*	*Z*	*E*	*Z*	*E*	*Z*	*E*	*Z*	*E*	*Z*	*E*	*Z*	(*Z*-*E*)
1	2.3617	2.1053	21.1967	20.0511	-16.9884	-12.7263	5.3041	2.6720	22.5889	22.9525	0.4099	1.3394	35.1113	36.5190	1.4077
2	2.2793	2.0998	1.1360	20.0920	-18.1119	-12.8479	5.0533	2.8018	23.0548	23.5098	0.5785	1.8802	34.2327	37.6771	3.4444
3	2.2635	1.7233	18.1455	19.9756	-14.0794	-7.0548	5.8469	3.2220	16.4471	16.2720	-0.8260	2.6942	28.0261	36.8613	8.8352
4	2.5255	2.0104	18.1607	20.0385	-20.3998	-13.4418	4.5750	2.0846	22.6106	22.5281	-0.7274	2.9618	26.9603	36.1932	9.2329
5	1.4867	1.6694	19.0825	20.3755	-15.4636	-15.1509	-3.2829	-0.4805	20.7267	20.6602	0.9071	-0.0928	23.6497	27.2394	3.5897
6	1.1908	1.1901	20.3437	20.3357	-15.8424	-15.8305	-2.0641	-2.0652	14.1684	14.1660	2.0822	2.0824	19.8509	19.8512	0.0003
7	2.9289	2.2214	23.6185	20.8137	-15.4076	-6.9241	6.1008	2.0226	21.5653	21.7513	-0.0592	0.7631	39.0216	40.7910	1.7694
8	2.0330	2.2415	20.2973	26.9058	-14.5679	-17.6490	-4.6619	-1.1011	19.5676	18.6052	1.4173	3.6445	24.0170	32.6400	8.623
9	1.5974	1.8407	16.9616	23.0831	-14.9627	-17.9503	-3.4649	0.0019	20.0960	19.1094	0.3474	2.5382	20.5614	28.6844	8.123
10	2.3730	2.1475	22.5495	21.4567	-11.8081	-7.4838	5.0765	2.5523	22.9717	23.3673	0.9222	1.8557	42.3556	44.0537	1.6981
11	1.4556	2.0559	19.9808	28.4685	-9.7094	-10.5460	-1.3111	-0.3999	19.2465	20.2425	1.4620	1.4181	31.3342	41.2880	9.9538
12	2.8699	2.2633	24.7224	22.2194	-9.8715	-1.6923	5.6004	1.9049	21.9160	22.1670	0.4251	1.1981	45.9653	48.2357	2.2704
13	2.8925	2.8741	29.8902	31.0545	-11.2512	-9.4784	3.3194	1.6203	18.4813	19.2414	-0.6517	0.2396	42.9079	45.6424	2.7345

Figure 5: Comparison of Total energy of *E* & *Z* isomers

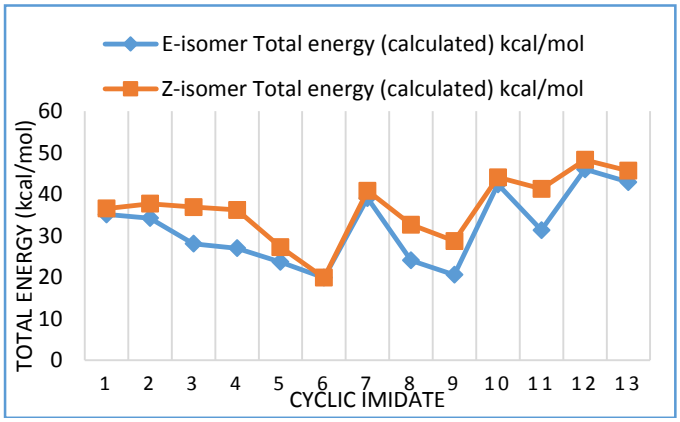
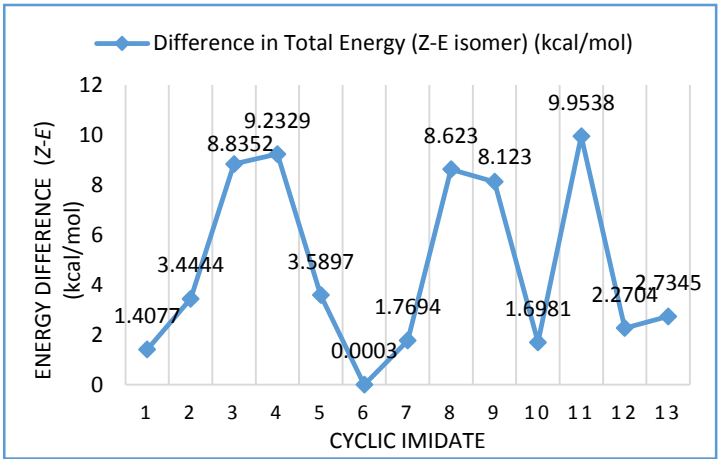


Figure 6: Energy difference between *E* & *Z* isomers



International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

EXPERIMENTAL SECTION

General procedure for calculation of total energy using molecular modelling experiments:

The desired compound structure was drawn in ChemDraw, and the structure was transferred to chem3D Pro 12.0. The desired molecular model could be drawn directly in chem3D Pro 12.0. A ball-and-stick molecular model appeared. Various models like wireframe, sticks, cylindrical bonds or space-filling models used also be used depending upon the size and structure of the molecule. The structure was cleaned up using the *Structure* tab. To obtain minimized energy values, the settings were kept to default; selected the *Calculations* toolbar, *Minimize Energy* job fuction was selected through the *MM2* tab. A small window displaying the minimum energy bar showed the four parameters namely *Job Type; Dynamics; Properties and General*. Under the *Job Type* bar, set the job type to *Minimize Energy* from the drop-down menu.

The columns displaying *Setup new Atom Types before Calculation*; *Setup new Atom Charges before Calculation* would be marked by default while the columns displaying *Display Every Iteration* and *Copy Measurements to Output Box* should be marked to get the complete information on the output tab. Set the value of *Minimum RMS Gradient* to 0.001. The *Run* button at the bottom of the Minimize Energy window is pressed. As the command is performed, an output tab appears at the bottom of the page where all the parameters (including stretch, bend, stretch-bend, torsion, Non-1,4 VDW, 1,4 VDW and Dipole/Dipole) along with the total steric energy (in kcal/mol) values were displayed. The results including the various parameters were recorded in their respective columns in the energy table (**Table 2**). For MM2 calculations of the other *Z*-geometric (not formed in experimental part) isomer structure was drawn in ChemDraw, copied and pasted through *Edit* tab (*Paste Special*) command from the submenu followed by the selection of *Preserve Coordinates.* After this, all other steps were repeated as defined in the procedure for the calculation of the total energy.

CONCLUSIONS

We believe that the MM2 molecular modelling calculations provide interesting results for the thermodynamic stability of the geometrical forms of cyclic imidates (**1-13**). The molecular modelling results were found to be in close agreement with the observed experimental results. This information may be useful for synthetic and medicinal chemists working on similar heterocyclic compounds, and may also help in developing new synthetic methodologies for similar compounds.

LIST OF ABBREVIATIONS

MM = Molecular Modelling
VDW = Van der Waals Forces
MMFF = Merck Molecular Force Field

REFERENCES

[1] Nevagi, R. J.; Dighe, S. N.; Dighe, S. N. Biological and Medicinal Significance of Benzofuran. *Eur. J. Med. Chem.* **2015**, *97*, 561–581.
[2] Royer, R. E.; Deck, L. M.; Vander Jagt, T. J.; Martinez, F. J.; Mills, R. G.; Young, S. A.; Vander Jagt, D. L. Synthesis and Anti-HIV Activity of 1,1'-Dideoxygossypol and Related Compounds. *J. Med. Chem.* **1995**, *38*, 2427–2432.
[3] Yu, Y.; Deck, J. A.; Hunsaker, L. A.; Deck, L. M.; Royer, R. E.; Goldberg, E.; Vander Jagt, D. L. Selective Active Site Inhibitors of Human Lactate Dehydrogenases A4, B4, and C4. *Biochem. Pharmacol.* **2001**, *62*, 81–89.
[4] Royer, R. E.; Mills, R. G.; Young, S. A.; Vander Jagt, D. L. Comparison of the Antiviral Activities of 3′-Azido-3′-Deoxythymidine (AZT) and Gossylic Iminolactone (GIL) against Clinical Isolates of HIV-1. *Pharmacol. Res.* **1995**, *31*, 49–52.
[5] Hegde, S. G.; Bryant, R. D.; Lee, L. F.; Parrish, S. K.; Parker, W. B. Cyclic Imidate Derivatives of 5-Amino-2,6-bis(polyfluoroalkyl)-pyridine-3-carboxylates, Synthesis and Herbicidal Activity. In *Synthesis and Chemistry of Agrochemicals IV*; ACS Symposium Series 584; American Chemical Society: Washington DC, 1995; Chapter 6, pp

60-69.
[6] Arunachalam, T.; Fan, H.; Pillai, K. M. R.; Ranganathan, R. S. Heterocyclic Nonionic X-Ray Contrast Agents. 4. The Synthesis of Dihydro-2(3H)-Furanylidenamino, 5-Oxo-1-pyrrolidinyl, 5-Oxo-4-Morpholinyl Derivatives by an Intramolecular Iodocyclization Approach. *J. Org. Chem.* **1995**, *60*, 4428–4438.
[7] Noël, T.; Bert, K.; Janssens, P.; Van der Eycken, J. Chiral Imidate Ligands: Synthesis and Applications in Asymmetric Catalysis. In *Innovative Catalysis in Organic Synthesis*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2012; pp 309–325.
[8] Xiong, T.; Zhang, Q.; Zhang, Z.; Liu, Q. A Divergent Synthesis of γ-Iminolactones, Dihydroquinolin-2-Ones, and γ-Lactames from β-Hydroxymethylcyclopropanylamides. *J. Org. Chem* **2007**, *72*, 8005–8009.
[9] Liu, G.; Zhou, Y.; Ye, D.; Zhang, D.; Ding, X.; Jiang, H.; Liu, H. Silver-Catalyzed Intramolecular Cyclization of o-(1-Alkynyl)benzamides: Efficient Synthesis of (1H)-Isochromen-1-Imines. *Adv. Synth. Catal.* **2009**, *351*, 2605–2610.
[10] Bian, M.; Yao, W.; Ding, H.; Ma, C. Highly Efficient Access to Iminoisocoumarins and α-Iminopyrones via AgOTf-Catalyzed Intramolecular Enyne−Amide Cyclization. *J. Org. Chem* **2010**, *75*, 269–272.
[11] Mehta, S.; Yao, T.; Larock, R. C. Regio- and Stereoselective Synthesis of Cyclic Imidates via Electrophilic Cyclization of 2-(1-Alkynyl)benzamides. A Correction. *J. Org. Chem* **2012**, *77*, 10938–10944.
[12] Mehta, S.; Waldo, J. P.; Neuenswander, B.; Lushington, G. H.; Larock, R. C. Solution-Phase Parallel Synthesis of a Multisubstituted Cyclic Imidate Library. *ACS Comb. Sci.* **2013**, *15*, 247–254.
[13] Schlemmer, C.; Andernach, L.; Schollmeyer, D.; Straub, B. F.; Opatz, T. Iodocyclization of o-Alkynylbenzamides Revisited: Formation of Isobenzofuran-1(3H)-Imines and 1H -Isochromen-1-imines Instead of Lactams. *J. Org. Chem.* **2012**, *77*, 10118–10124.
[14] Trabulsi, H.; Guillot, R.; Rousseau, G. Preparation of Imino Lactones by Electrophilic Cyclization of β,γ-Unsaturated Hydroxamates: Formation of 3-Cyanoprop-2-en-1-ones through Fragmentation Reactions. *Eur. J. Org. Chem.* **2010**, 5884–5896.
[15] Davidson, E. R. Molecular Mechanics and Modeling: Overview. *Chem. Rev.* **1993**, *93*, 2337–2337.
[16] Lipkowitz, K. B.; Peterson, M. A. Molecular Mechanics in Organic Synthesis. *Chem. Rev.* **1993**, *93*, 2463–2486.
[17] Parrill, A. L., Lipkowitz, K. B., Eds.; Reviews in Computational Chemistry; John Wiley & Sons, Inc: Hoboken, NJ, **2016**.
[18] Dudek, M. J.; Ponder, J. W. Accurate Modeling of the Intramolecular Electrostatic Energy of Proteins. *J. Comput. Chem.* **1995**, *16*, 791–816.
[19] Burkert, U; Allinger, N.L. Molecular Mechanics. *ACS Monograph 177*, American Chemical Society, Washington, DC (1982); **1982**.
[20] J. Andrew McCammon, S. C. H. Dynamics of Proteins and Nucleic Acids, 1988th ed.; Cambridge University Press, **1988**.
[21] Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94*. *J. Comput. Chem.* **1996**, *17*, 490–519.
[22] Lii, J. H.; Allinger, N. L. Molecular Mechanics. The MM3 Force Field for Hydrocarbons. 3. The van der Waals' Potentials and Crystal Data for Aliphatic and Aromatic Hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8576–8582.
[23] Allinger, N. L. Understanding Molecular Structure from Molecular Mechanics. *J Comput Aided Mol Des* **2011**, *25*, 295–316.
[24] Ren, P.; Ponder, J. W. Consistent Treatment of Inter- and Intramolecular Polarization in Molecular Mechanics Calculations. *J. Comput. Chem.* **2002**, *23*, 1497–1506.
[25] Schnur, D. M.; Grieshaber, M. V; Bowen, J. P. Development of an Internal Searching Algorithm for Parameterization of the MM2/MM3 Force Fields. *J. Comput. Chem.* **1991**, *12*, 844–849.
[26] Vanommeslaeghe, K.; Guvench, O.; MacKerell, A. D. Molecular Mechanics. *Curr Pharm Des.* **2014**, *20*, 3281–3292.

arXiv:1805.08172v1 [quant-ph] 21 May 2018

Proposal for Dimensionality Testing in Quantum Private Query

Arpita Maitra
Centre for Theoretical Studies
Indian Institute of Technology Kharagpur

Bibhas Adhikari
Department of Mathematics
Indian Institute of Technology Kharagpur

Satyabrata Adhikari
Department of Applied Mathematics
Delhi Technological University

Recently, dimensionality testing of a quantum state has received extensive attention (Acín et al. Phys. Rev. Letts. 2006, Scarani et al. Phys. Rev. Letts. 2006). Security proofs of existing quantum information processing protocols rely on the assumption about the dimension of quantum states in which logical bits will be encoded. However, removing such assumption may cause security loophole. In the present report, we show that this is indeed the case. We choose two player's quantum private query protocol by Yang et al. (Quant. Inf. Process. 2014) as an example and show how one player can gain considerable information than desired by changing the dimension of subsystem of a shared quantum system. To resist such attack we propose dimensionality testing in a different way. Our proposal is based on CHSH like game which certifies the degrees of freedom of the subsystems. As we exploit CHSH like game, it can be used to test if the states are separable for which the protocol becomes completely vulnerable.

I. INTRODUCTION

Testing the dimension of a quantum state has generated a lot of interest recently due to its application in ascertaining security of various quantum information processing protocols [1, 2]. In[1] and [3] it was questioned that what if the assumption about the dimension of quantum systems in which logical bits are encoded will be removed. Thus it has redirected the thoughts to derive bounds for the security proofs of quantum information processing protocols on weaker constraints. The use of higher/lower dimensional state for quantum information retrieval may provide a party in terms of probability to know certain information than desired. In particular, for a two-party information sharing protocol using shared randomness, the change of dimension of the state of one of the shared particles may influence information gain for a party involved in the dealing.

In this direction, detection of the dimension of an unknown quantum system based on a set of the conditional probabilities of observing an outcome after creating and measuring the system for a given set of possibilities have become a prominent research area [4–6]. Successful experimental tests are also carried out for testing dimension of a quantum system [7, 8]. However, these attempts to decide about the dimension of a state are proposed in a prepare-measurement set up with/without the aid of dimension witnesses [9, 10].

In this paper we develop a CHSH like game which helps in determining the degrees of freedom of one subsystem of an entangled bipartite system. Indeed, we consider an unknown shared entangled state of the form

$$|\psi\rangle_{BA} = \frac{1}{\sqrt{2}}(|0\rangle_B |\phi_0\rangle_A + |1\rangle_B |\phi_1\rangle_A), \qquad (1)$$

which is shared between two legitimate parties Bob and Alice and $\langle\phi_0|\phi_1\rangle_A \neq 0$. In general, $\{|0\rangle_B, |1\rangle_B\}$ denotes the computational basis for Bob's qubits and $|\phi_l\rangle_A, l = 0, 1$ denotes qutrits with two degrees of freedom. Note that, given an orthonormal basis, say $\{|0\rangle_A, |1\rangle_A, |2\rangle_A\}$ of the 3 dimensional Hilbert space $\mathbb{C}^3$, by two degrees of freedom of a qutrit we mean if the qutrit $|\phi\rangle_l$ is in the span of $\{|i\rangle_A, |j\rangle_A\}, i, j \in \{0, 1, 2\}$, that is, the state is in the superposition of any two basis vectors out of three. Whereas, $|\phi\rangle_l$ has three degrees of freedom if it is a superposition of all the three orthonormal basis vectors. The subscripts A and B stand for Alice and Bob respectively. The dimension testing problem which we consider in this paper is concerned with the fact that Bob has to determine whether both $|\phi_0\rangle_A, |\phi_1\rangle_A$ are having two degrees of freedom lying in the same subspace of $\mathbb{C}^3$ or they lie in different subspaces of $\mathbb{C}^3$. We solve the problem by defining a CHSH like game. The proposed game is based on a function which generally familiar as embedded XOR function [11]. We calculate the winning probability of the game for separable and entangled states. We notice that the winning probability of the game differs for separable state from entangled one. We also notice that if the entangled particles are not in the same Hilbert space then the winning probability changes abruptly. Observing this success probability one can certify if the states are in desired form. The advantage of the proposed game for dimension testing is that it can be employed when the entangled states used by legitimate

parties are borrowed from a third party and the dimension testing is possible without preparation-measurement framework. This degree of freedom detection problem is motivated by the following reason. Many quantum information retrieval protocols between two legitimate parties are based on shared entangled states of the form (1). Such a protocol typically starts with sending out a sequence of subsystems of the bipartite systems from Bob to Alice that will be used for generation of a secret key using local measurements. After sending the states to Alice, Bob measures his qubits that are entangled with Alice's states sequentially in $\{|0\rangle_B, |1\rangle_B\}$ basis, whereas Alice measures her qubits either in $\{|\phi_0\rangle_A, |\phi_0^{\perp}\rangle_A\}$ basis or in $\{|\phi_1\rangle_A, |\phi_1^{\perp}\rangle_A\}$ basis randomly. If the measurement result of Alice gives $|\phi_0^{\perp}\rangle$, she concludes that the raw key bit at Bob's end must be 1. If it is $|\phi_1^{\perp}\rangle$, the raw key bit must be 0. Bob and Alice execute classical post-processing so that Alice's information on the key reduces to one bit or more. Bob knows the whole key, whereas Alice generally knows several bits of the key. For example, in the quantum private query protocol due to Yang et al. [12], set $|\phi_0\rangle_A = \cos(\frac{\theta}{2})|0\rangle + \sin(\frac{\theta}{2})|1\rangle$, $|\phi_1\rangle_A = \cos(\frac{\theta}{2})|0\rangle - \sin(\frac{\theta}{2})|1\rangle$, $0 < \theta < \pi/2$. Then it can be shown that the success probability of Alice to guess a bit in the raw key is $\frac{1}{2}\sin^2\theta$. Indeed, if Bob has lack of resources to create such states $|\psi\rangle_{AB}$ and he borrows the states from a third party, say Charlie, then the situation would be different. In fact, if Alice is mistrustful and has a tie with Charlie, then there may exist a possibility that the states $|\phi_l\rangle_A, l = 0, 1$ are not qubits. Rather, the states $|\phi_l\rangle_A$ may be of higher dimensional which may benefit Alice in terms of success probability of knowing some bits in the raw key. We show in this paper that this is indeed possible and thus the key generation in quantum private query protocol (QPQ) proposed by Yang et al. is insecure without the certification of dimension of the Alice's states that are shared and entangled with Bob. Besides, we show by defining a qubit-qutrit entangled state $|\Psi\rangle_{BA}$ that if this state is supplied by Charlie to Bob then it yields higher success probability to Alice to guess a bit in the raw key for the QPQ protocol proposed by Yang et al. We also demonstrate that it is possible to create such an entangled state $|\Psi\rangle_{BA}$ in an experimental setup. This observation demands the justification and usefulness for the need to test the dimension of Alice's state in the shared entangled state before starting the protocol. In order to acquire knowledge about the dimension of Alice's particle, Bob needs to perform certain quantum measurements. Indeed, since the dimension of Alice's state is unknown to Bob and he has to devise measurement operators for detecting whether it is a qubit or qutrit, we define a set of entangled states $|\Psi\rangle_{BA}$ such that Alice's state is a qutrit with two degrees of freedom and enables the success probability equal to the same while using the qubits. Thus, when Yang et al. protocol is used for such states, it maintains the desired success probability even if a qubit-qutrit entangled state is used between them and Bob can apply his measurement operators for detection of the degrees of freedom of the qutrits shared with Alice.

Moreover, the procedure we exploit for dimension certification in the above mentioned protocol can detect a more powerful attack as follows. In this attack model Charlie may supply n product states of the form $|l\rangle_B |\phi_l\rangle_A$, where $l \in \{0, 1\}$, to Bob. At the same time he provides the full information of l to Alice. As Bob measures his states only in $\{|0\rangle, |1\rangle\}$ basis, Alice gets the full information about the raw key. Though such type of attack was not considered in [13], however, the methodology they used certifies automatically if the states are entangled and hence remove the possibility of such attack. We show that the CHSH like game proposed in the paper is capable of defending such an attack.

The contribution and organization of this paper are as follows. In Section II, we develop a CHSH like game for detection of dimension of Alice's subsystem for a shared entangled state shared between Bob and Alice of the form (1). In Section III, we define a qubit-qutrit entangled state $|\Psi\rangle_{BA}$ which guarantees higher success probability to Alice to guess a bit in the raw key for the QPQ protocol proposed by Yang et al. We also propose an experimental set up for the design of such a state. Next, we define a set of entangled qubit-qutrit states $|\Psi\rangle_{BA}$ such that Alice's state is with two degrees of freedom and it enables the success probability equal to the same while using the qubits for the Yang et al. protocol.

II. CHSH LIKE GAME FOR DIMENSIONALITY TESTING

Alice's qutrits can have two possible degrees of freedom, that is, 2 and 3 ($0 < \theta < \pi/2$). In our context, Bob performs a CHSH like game to detect whether the states of Alice's particles, $|\phi_0\rangle_A, |\phi_1\rangle_A$ (see equation (1)) have two degrees of freedom that are in the superposition of same orthonormal states $\{|i\rangle_A, |j\rangle_A\}, i, j \in \{0, 1, 2\}$, or the shared states are separable states. Similar to the standard CHSH game, we require two black boxes as initial set up. One box is labeled as A whereas another is labeled as B. A can take an input $x \in \{0, 1\}$ and B can take another input $y \in \{0, 1\}$. After taking the inputs, A produces a bit $a \in \{0, 1\}$ and B produces a trit $b \in \{0, 1, 2\}$. We now define a function $f(a, b)$ such that $f(a, b) = 1$ if $a \neq b$ and $f(a, b) = 0$ if $a = b$.

Now, we consider two players $\mathbb{A}$ and $\mathbb{B}$ who are holding boxes A and B respectively. The players win the game if and only if $f(a, b) = x \wedge y$. For classical deterministic strategy the winning probability of the game is $\frac{3}{4}$. However, the winning probability differs if we assume that the players share some quantum states between themselves. In the following subsections we will show how it differs for separable state, entangled state with sub-systems lying in same subspaces of dimension 2 and entangled state with sub-systems lying in different subspaces of dimension 2.

A. Winning probability for product states

Let the players $\mathbb{A}$ and $\mathbb{B}$ share n separable states between themselves. Let among these n states, $\frac{n}{2}$ states are of the form $|0\rangle|\phi_0\rangle$ and remaining $\frac{n}{2}$ are of the form $|1\rangle|\phi_1\rangle$, where $|\phi_0\rangle_A = \cos\frac{\theta}{2}|i\rangle + \sin\frac{\theta}{2}|i+1\rangle$ and $|\phi_1\rangle_A = \cos\frac{\theta}{2}|i\rangle - \sin\frac{\theta}{2}|i+1\rangle$, $i \in \{0,1,2\}$ and $|i+1\rangle = |i+1 \bmod 3\rangle$. Player $\mathbb{A}$ holds box A whereas player $\mathbb{B}$ holds box B. Now, the players fix a quantum strategy as follows.

If $x = 0$, A measures the particle in $\{|0\rangle, |1\rangle\}$ basis, if it is 1, the particle is measured in $\{|+\rangle, |-\rangle\}$ basis, where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. If the measurement result would be $|0\rangle$ or $|+\rangle$, A outputs 0. If the measurement result would be $|1\rangle$ or $|-\rangle$, A outputs 1.

If $y = 0$, B measures the particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, if it is 1, the particle is measured in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis, where $|0'\rangle = \frac{1}{\sqrt{2}}(\cos\frac{\pi}{8}|i\rangle + \sin\frac{\pi}{8}|i+1\rangle)$, $|1'\rangle = \frac{1}{\sqrt{2}}(\sin\frac{\pi}{8}|i\rangle - \cos\frac{\pi}{8}|i+1\rangle)$, $|2'\rangle = |i+2\rangle$ and $|0''\rangle = \frac{1}{\sqrt{2}}(\cos\frac{3\pi}{8}|i\rangle + \sin\frac{3\pi}{8}|i+1\rangle)$, $|1''\rangle = \frac{1}{\sqrt{2}}(\sin\frac{3\pi}{8}|i\rangle - \cos\frac{3\pi}{8}|i+1\rangle)$, $|2''\rangle = |i+2\rangle$; $|i+2\rangle = |i+2 \mod 3\rangle$. If the measurement result would be $|0'\rangle$ or $|0''\rangle$, B outputs 0. If the measurement result would be $|1'\rangle$ or $|1''\rangle$, B outputs 1. If it is $|2'\rangle$ or $|2''\rangle$, B outputs 2.

Thus the winning probability in this case becomes $p(f(a,b) = x \wedge y) = p((x,y) = (0,0)\ \&\ ((a,b) = (0,0)\ or\ (1,1))) + p((x,y) = (0,1)\ \&\ ((a,b) = (0,0)\ or\ (1,1))) + p((x,y) = (1,0)\ \&\ ((a,b) = (0,0)\ or\ (1,1))) + p((x,y) = (1,1)\ \&\ ((a,b) = (0,1)\ or\ (0,2)\ or\ (1,0)\ or\ (1,2)))$. From table I of supplementary material [14], we get $p(f(a,b) = x \wedge y) = \frac{1}{2}(1 + \frac{1}{2\sqrt{2}}\sin\theta)$.

B. Winning probability for entangled state with sub-systems lying in same subspaces

Let $\mathbb{A}$ and $\mathbb{B}$ share n entangled pairs of the form $\frac{1}{\sqrt{2}}(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle)$, where $|\phi_0\rangle_A = \cos\frac{\theta}{2}|i\rangle + \sin\frac{\theta}{2}|i+1\rangle$ and $|\phi_1\rangle_A = \cos\frac{\theta}{2}|i\rangle - \sin\frac{\theta}{2}|i+1\rangle$. They follow the quantum strategy as described above. In such a case, the winning probability $p(f(a,b) = x \wedge y)$ becomes $\frac{1}{2}(1 + \frac{1}{2\sqrt{2}} + \frac{1}{2\sqrt{2}}\sin\theta)$ (table II in supplementary material).

C. Winning probability for entangled states with subsystem lying in different subspaces

Let $\mathbb{A}$ and $\mathbb{B}$ share n entangled pairs of the form $\frac{1}{\sqrt{2}}(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle)$, where $|\phi_0\rangle = \cos\theta|i+1\rangle + \sin\theta|i+2\rangle$ and $|\phi_1\rangle = \cos\theta|i\rangle - \sin\theta|i+2\rangle$. Now, let the players follow the quantum strategy described above. In such case, the winning probability $p(f(a,b) = x \wedge y)$ becomes $\frac{1}{4}(1 + \cos^2\theta)$ (table III in supplementary material).

Thus, observing the winning probability of the game we can differentiate if the states are separable, entangled with sub-systems lying in same subspaces or entangled with sub-systems lying in different subspaces. Figure 1 shows the winning probabilities with varying θ for the above three cases.

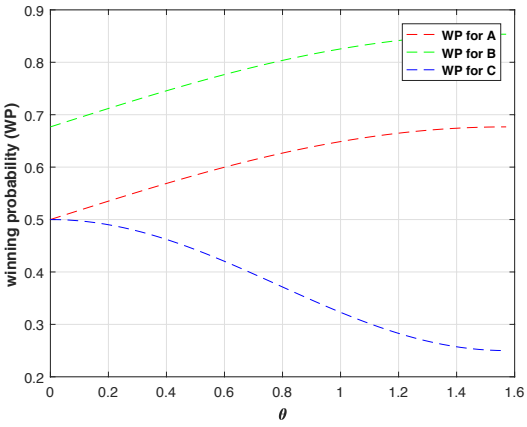


FIG. 1: A: Winning probability (WP) for tilted CHSH game with θ for separable states, B: WP for entangled state with sub-systems lying in same subspace, C: WP for entangled state with sub-systems lying in different subspace

III. PROPOSAL AND GENERATION OF A QUBIT-QUTRIT SHARED ENTANGLED STATE

In this section we show how the success probability for guessing a raw key bit in a typical key generation protocol using shared entangled qubit-qubit state may get influenced by replacing it with a qubit-qutrit shared entangled state. We also demonstrate a procedure for creating such qubit-qutrit entangled states in an experimental set up using standard quantum logic gates.

Define

$$|\Psi\rangle_{BA} = \frac{1}{\sqrt{2}}(|0\rangle_B|\phi_0\rangle_A + |1\rangle_B|\phi_1\rangle_A) \tag{2}$$

where $|\phi_0\rangle_A = \cos\gamma\cos\delta|i\rangle + (\cos\theta\sin\delta - \sin\theta\sin\gamma\cos\delta)|i+1\rangle + (\sin\theta\sin\delta + \cos\theta\sin\gamma\cos\delta)|i+2\rangle$, $|\phi_1\rangle_A = (\cos\theta\sin\delta - \sin\theta\sin\gamma\cos\delta)|i\rangle + \cos\gamma\cos\delta|i+1\rangle - (\sin\theta\sin\delta + \cos\theta\sin\gamma\cos\delta)|i+2\rangle$, $i \in \{0,1,2\}$, $0 \le \theta, \gamma, \delta \le \pi/2$. Note that $|\phi_0\rangle_A$ and $|\phi_1\rangle_A$ need not be orthogonal.

Now we discuss a standard key generation protocol, for example Yang et al. protocol [12] using this shared qubit-qutrit entangled state. After sharing the states, Bob measures his qubits in $\{|0\rangle_B, |1\rangle_B\}$ basis, whereas Alice measures her qutrits either in $\{|\phi_0\rangle_A, |\phi_0'\rangle_A, |\phi_0''\rangle_A\}$ basis or in

$\{|\phi_1\rangle_A, |\phi_1'\rangle_A, |\phi_1''\rangle_A\}$ basis randomly, where $|\phi_0'\rangle_A = -\cos\gamma\sin\delta\,|i\rangle + (\sin\theta\sin\gamma\sin\delta + \cos\theta\cos\delta)\,|i+1\rangle + (\sin\theta\cos\delta - \sin\delta\cos\theta\sin\gamma)\,|i+2\rangle, |\phi_0''\rangle_A = -\sin\gamma\,|i\rangle - \sin\theta\cos\gamma\,|i+1\rangle + \cos\theta\cos\gamma\,|i+2\rangle, |\phi_1'\rangle_A = (\sin\theta\sin\gamma\sin\delta + \cos\theta\cos\delta)\,|i\rangle - \cos\gamma\sin\delta\,|i+1\rangle - (\sin\theta\cos\delta - \sin\delta\cos\theta\sin\gamma)\,|i+2\rangle, |\phi_1''\rangle_A = -\sin\theta\cos\gamma\,|i\rangle - \sin\gamma\,|i+1\rangle - \cos\theta\cos\gamma\,|i+2\rangle$.

If the measurement outcome of Alice is $|\phi_0'\rangle$ or $|\phi_0''\rangle$, she concludes that the raw key bit at Bob's end is 1. If it is $|\phi_1'\rangle$ or $|\phi_1''\rangle$, the raw key bit is 0. Then the success probability of Alice when Bob measures $|0\rangle$ is as follows. $p(A=0,B=0) = p(B=0)p(A=0|B=0) = \frac{1}{2}[p(A=\phi_1'|B=0) + p(A=\phi_1''|B=0)] = \frac{1}{2}\big[(\sin\theta\sin\gamma\cos\gamma\sin 2\delta + \cos\theta\cos\gamma\cos 2\delta - \sin\theta\cos\theta\sin\gamma\cos 2\delta - \sin^2\theta\sin\delta\cos\delta + \cos^2\theta\sin^2\gamma\sin\delta\cos\delta)^2 + (-\sin\theta\cos\delta\cos 2\gamma - \cos\theta\sin\gamma\sin\delta - \sin\theta\cos\theta\cos\gamma\sin\delta - \cos^2\theta\sin\gamma\cos\gamma\cos\delta)^2\big]$.

Similarly, the success probability of Alice when Bob measures $|1\rangle$ is given by $p(A=1,B=1) = \frac{1}{2}\big[(\sin\theta\sin\gamma\cos\gamma\sin 2\delta + \cos\theta\cos\gamma\cos 2\delta - \sin\theta\cos\theta\sin\gamma\cos 2\delta - \sin^2\theta\sin\delta\cos\delta + \cos^2\theta\sin^2\gamma\sin\delta\cos\delta)^2 + (-\sin\theta\cos\delta\cos 2\gamma - \cos\theta\sin\gamma\sin\delta - \sin\theta\cos\theta\cos\gamma\sin\delta - \cos^2\theta\sin\gamma\cos\gamma\cos\delta)^2\big]$.

Hence the total success probability of Alice to guess a bit correctly can be calculated as follows $p(A=B) = \frac{1}{2}.[p(A=\phi_1'|B=0) + p(A=\phi_1''|B=0)] + \frac{1}{2}.[p(A=\phi_0'|B=1) + p(A=\phi_0''|B=1)] = [p(A=\phi_0'|B=1) + p(A=\phi_0''|B=1)] = (\sin\theta\sin\gamma\cos\gamma\sin 2\delta + \cos\theta\cos\gamma\cos 2\delta - \sin\theta\cos\theta\sin\gamma\cos 2\delta - \sin^2\theta\sin\delta\cos\delta + \cos^2\theta\sin^2\gamma\sin\delta\cos\delta)^2 + (-\sin\theta\cos\delta\cos 2\gamma - \cos\theta\sin\gamma\sin\delta - \sin\theta\cos\theta\cos\gamma\sin\delta - \cos^2\theta\sin\gamma\cos\gamma\cos\delta)^2$.

Observe that setting $\delta = \pi/2$, the success probability of Alice to guess a key bit correctly becomes $\cos^2\theta(1+\sin^2\theta) = 1-\sin^4\theta$ for any $0 \le \theta, \gamma \le \pi/2$. The success probabilities for Yang et al. protocol (see Section II) are drwan in Figure 1 for both the qubit-qubit (given by eqn (1)) and qubit-qutrit (given by eqn (2), $\delta = \pi/2$) shared entangled states. Note that the success probability of using the shared state as (2), Alice gains in terms of probability for any value of θ ranging from 0 to 1.1 (approximately), and hence this observation demands Bob to verify and test the dimension of the quantum particle shared with Alice.

However, we now show that even if Alice's subsystems are qutrits the success probability of the key generation protocol remains same as the success probability with qubit-qubit system. We propose the following qubit-qutrit entangled state for the same.

Define

$$|\Phi\rangle_{BA} = \frac{1}{\sqrt{2}}(|0\rangle_B|\phi_0\rangle_A + |1\rangle_B|\phi_1\rangle_A) \qquad (3)$$

where $|\phi_0\rangle_A = \cos\frac{\theta}{2}|i\rangle + \sin\frac{\theta}{2}|i+1\rangle$ and $|\phi_1\rangle_A = \cos\frac{\theta}{2}|i\rangle - \sin\frac{\theta}{2}|i+1\rangle, 0 < \theta < \pi/2, i \in \{0,1,2\}$. Note

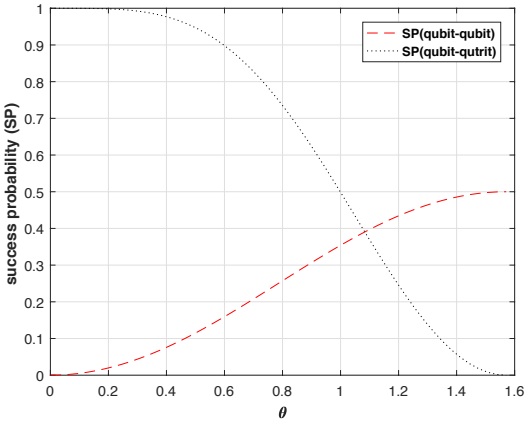


FIG. 2: Success probability for Yang's QPQ protocol for qubit-qubit and qubit-qutrit (setting $\delta = \pi/2$) shared entangled state

that if Bob's particle is measured in $\{|0\rangle_B, |1\rangle_B\}$ basis then the state of Alice's particle lies in one of the fundamental two dimensional subspaces of $\mathbb{C}^3$. Now we determine the success probability of Alice to guess a bit of the raw key as follows.

Proceeding a similar way described by Yang et al. [12], first Bob measures his qubits in $\{|0\rangle_B, |1\rangle_B\}$ basis. If Bob obtains $|0\rangle_B$ or $|1\rangle_B$, Alice's state becomes $|\phi_0\rangle_A$ or $|\phi_1\rangle_A$ respectively. Now let Alice perform measurements on her particle using the bases $\mathcal{A}_0 = \{|\phi_0\rangle_A, |\phi_0'\rangle_A, |\phi_0''\rangle_A\}$ and $\mathcal{A}_1 = \{|\phi_1\rangle_A, |\phi_1'\rangle_A, |\phi_1''\rangle_A\}$, choosing one of them uniformly at random where $|\phi_l'\rangle_A$ is in the superposition of $|i\rangle, |i+1\rangle$ and orthogonal to $|\phi_l\rangle_A$, and $|\phi_l''\rangle_A$ is a unit vector orthogonal to both $|i\rangle$ and $|i+1\rangle, l = 0,1$. If Bob obtains $|0\rangle_B$ and Alice chooses $\mathcal{A}_0$, she shall get $|\phi_0\rangle_A$ with probability 1 and never gets $|\phi_0'\rangle_A, |\phi_0''\rangle_A$; whereas if she chooses $\mathcal{A}_1$, she shall obtain $|\phi_1\rangle_A$ with probability $\cos^2\theta$ otherwise $|\phi_1'\rangle_A$ with probability $\sin^2\theta$ and never gets $|\phi_1''\rangle_A$. Indeed, note that $|\phi_0\rangle_A = \cos\frac{\theta}{2}\left(\cos\frac{\theta}{2}|\phi_1\rangle_A + \sin\frac{\theta}{2}|\phi_1'\rangle_A\right) + \sin\frac{\theta}{2}\left(-\sin\frac{\theta}{2}|\phi_1\rangle_A + \cos\frac{\theta}{2}|\phi_1'\rangle_A\right) = \cos\theta|\phi_1\rangle_A + \sin\theta|\phi_1'\rangle_A$. Thus we summerize the conditional probabilities in the following table, where $B=0,1$ mean Bob gets $|0\rangle_B$ and $|1\rangle_B$ respectively.

	$B=0$	$B=1$
$A=\|\phi_0\rangle_A$	$\frac{1}{2}$	$\frac{1}{2}\cos^2\theta$
$A=\|\phi_0'\rangle_A$	0	$\frac{1}{2}\sin^2\theta$
$A=\|\phi_0''\rangle_A$	0	0
$A=\|\phi_1\rangle_A$	$\frac{1}{2}\cos^2\theta$	$\frac{1}{2}$
$A=\|\phi_1'\rangle_A$	$\frac{1}{2}\sin^2\theta$	0
$A=\|\phi_1''\rangle_A$	0	0

We define the rule to determine the key as follows. If Alice gets $|\phi_0'\rangle_A$, she outputs 1, and when she gets $|\phi_1'\rangle_A$, she outputs 0. Thus, the success probability of Alice to guess a bit in raw key can be written as $p(A=B) =$

$p(A=0,B=0)+p(A=1,B=1)=p(B=0)p(A=0|B=0)+p(B=1)p(A=1|B=1)=\frac{1}{2}p(A=\phi_1^{\perp}|B=0)+\frac{1}{2}p(A=\phi_0^{\perp}|B=1)=\frac{\sin^2\theta}{2}$.

Thus we conclude that the proposed qubit-qutrit state $|\Phi\rangle_{BA}$ provides the same success probability when the state of Alice's shared particle is in one of the fundamental subspaces of $\mathbb{C}^3$ after Bob's measurement of his qubit. This result facilitates us to define a set of measurement operators for Bob who can test whether Alice's qutrit is in the desired space. Once a dimension test determine that Alice's particle is having two degrees of freedom, that is, lying in the same two dimensional subspace of $\mathbb{C}^3$ independent of Bob's measurement outcome, the Yang et al. protocol can be followed for key generation with the shared entangled state of the form eqn (3).

Now we explain how is it possible to create a pair of entangled particles defined by the state in eqn (2) in an experimental setup.

IV. EXPERIMENTAL DEMONSTRATION OF QUBIT-QUTRIT ENTANGLED SYSTEM

In this section, we propose an experimental set-up to generate qubit-qutrit entangled pair which is important to implement our proposal experimentally.

Assume that both the qubit and qutrit are initially in a vacuum mode and the initial qubit-qutrit state given by $|\mu\rangle_{BA}=|0\rangle_B\otimes|0\rangle_A$. Our goal is to use different quantum gates to generate entanglement in qubit-qutrit system which is initially in a product state. Recall that the rotation operator R for a qutrit system can be written as $R=R_x(\theta)R_y(\gamma)R_z(\delta)$ due to Euler decomposition where $R_x(\theta)$, $R_y(\gamma)$ and $R_z(\delta)$ denote the rotation operators about x-axis, y-axis and z-axis respectively. Besides, the operators $R_x(\theta)$, $R_y(\gamma)$ and $R_z(\delta)$ can be realized in experiment by optical elements such as beam splitters and a π-phase shifter.

Now define a unitary operator U which acts on the computational basis state of six dimensional Hilbert space such that $U|00\rangle=|00\rangle, U|01\rangle=|01\rangle, U|02\rangle=|02\rangle, U|10\rangle=|11\rangle, U|11\rangle=|10\rangle, U|12\rangle=-|12\rangle$. Therefore, the explicit form of the unitary operator is given by $U=|00\rangle\langle00|+|01\rangle\langle01|+|11\rangle\langle10|+|10\rangle\langle11|+|02\rangle\langle02|-|12\rangle\langle12|$.

Then the desired state of the form (2) can be obtained from $|\mu\rangle_{BA}$ in two steps as follows. In the first step, apply the unitary operator $H\otimes R$ on $|\mu\rangle_{BA}$ such that we obtain $|\omega\rangle_{BA}=(H\otimes R)|\mu\rangle_{BA}$, where H denotes the Hadamard gate for qubits. Then apply the operator U to obtain $|\Psi\rangle_{BA}=U|\omega\rangle_{BA}$. Indeed, writing R in the computation basis, it is easy to verify that $|\Psi\rangle_{BA}=\frac{1}{\sqrt{2}}(|0\rangle_B\otimes|\phi_0\rangle_A+|1\rangle_B\otimes|\phi_1\rangle_A)$, where $|\phi_0\rangle_A=\cos(\gamma)\cos(\delta)|0\rangle_A+(\cos(\theta)\sin(\delta)-\sin(\theta)\sin(\gamma)\cos(\delta))|1\rangle_A+(\sin(\theta)\sin(\delta)+\cos(\theta)\sin(\gamma)\cos(\delta))|2\rangle_A$ and $|\phi_1\rangle_A=(\cos(\theta)\sin(\delta)-\sin(\theta)\sin(\gamma)\cos(\delta))|0\rangle_A+\cos(\gamma)\cos(\delta)|1\rangle_A-(\sin(\theta)\sin(\delta)+\cos(\theta)\sin(\gamma)\cos(\delta))|2\rangle_A$.

V. CONCLUSION

All the existing quantum information processing protocols assume a certain dimension of the system. It is intuitively commented that removing such assumption may cause any security flaw in quantum information tasking. However, till date, no such protocol is found which can prove this conjecture. In the present draft, we find that there exist at least one key generation protocol which suffers from the removal of such assumption. In this regard, we pick up quantum private query protocol by Yang et al. and show how one party can gain more information than suggested by the protocol by changing the dimension of his/her subsystem. In this initiative, we propose titled CHSH game to certify the dimension of the subsystems of shared quantum system. Along with the certification of dimensionality the game is enable to certify if the states are entangled.

We propose an experimental set-up to generate qubit-qutrit entangled pair which is the backbone of our current proposal.

[1] Acín, Antonio and Gisin, Nicolas and Masanes, Lluis, *Physical Review Letters*, **97**, 12, 120405, 2006

[2] Brunner, Nicolas and Pironio, Stefano and Acín, Antonio and Gisin, Nicolas and Méthot, André Allan and Scarani, Valerio, *Physical Review Letters*, **100**, 21, 210503, 2008

[3] Scarani, Valerio and Gisin, Nicolas and Brunner, Nicolas and Masanes, Lluis and Pino, Sergi and Acín, Antonio, *Physical Review A*, **74**, 4, 042339, 2006

[4] Wehner, Stephanie and Christandl, Matthias and Doherty, Andrew C, *Physical Review A*, **78**, 6, 062112, 2008

[5] Gallego, Rodrigo and Brunner, Nicolas and Hadley, Christopher and Acín, Antonio, *Physical Review Letters*, **105**, 23, 230501, 2010

[6] Junge, Marius and Palazuelos, Carlos, *Communications in Mathematical Physics*, **306**, 3, 695–746, 2011

[7] Hendrych, Martin and Gallego, Rodrigo and Mičuda, Michal and Brunner, Nicolas and Acín, Antonio and Torres, Juan P, *Nature Physics*, **8**, 8, 588–591, 2012

[8] Ahrens, Johan and Badzig, Piotr and Cabello, Adán and Bourennane, Mohamed, *Nature Physics*, **8**, 8, 592–595, 2012

[9] Brunner, Nicolas and Navascués, Miguel and Vértesi, Tamás, *Physical Review Letters*, **110**, 15, 150501, 2013

[10] Bowles, Joseph and Quintino, Marco Túlio and Brunner, Nicolas, *Physical Review Letters*, **112**, 14, 140407, 2014

[11] Dov Gordon, S and Hazay, Carmit, and Katz, Jonathan and Lindell, Yehuda, *Journal of the ACM (JACM)*, **58**, 6, 24:1–24:37, 2011

[12] Yang, Yu-Guang and Sun, Si-Jia and Xu, Peng and Tian, Ju, *Quantum Information Processing*, **13**, 3, 805–813,

2014
[13] Maitra, Arpita and Paul, Goutam and Roy, Sarbani, *Physical Review A*, **95**, 4, 042344, 2017
[14] Supplementary Material

VI. APPENDIX

A. Conditional probability for separable state

Let us consider a situation where Charlie supplies $\frac{n}{2}$ separable states of the form $|0\rangle|\phi_0\rangle$ and $\frac{n}{2}$ separable states of the form $|1\rangle|\phi_1\rangle$, where $|\phi_0\rangle = \cos\frac{\theta}{2}|i\rangle + \sin\frac{\theta}{2}|i+1\rangle$ and $|\phi_1\rangle = \cos\frac{\theta}{2}|i\rangle - \sin\frac{\theta}{2}|i+1\rangle$. Now, we analyze the case by case scenario below.

1. *Case 1: (x=0, y=0)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. For $\frac{1}{2}$ of the cases, for the first particle he obtains $|0\rangle$ with probability 1 and in that case the second particle will be $|\phi_0\rangle$. When he measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, he obtains $|0'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8} + \sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$, $|1'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8} - \sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$. In such case, he never gets $|2'\rangle$. Similarly, for $\frac{1}{2}$ of the cases, Bob gets $|1\rangle$ with probability 1 and in that case the second particle will be $|\phi_1\rangle$. When he measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, he obtains $|0'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8} - \sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$, $|1'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8} + \sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$ and never gets $|2'\rangle$.

2. *Case 2: (x=0, y=1)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 1. Conditional probabilities $\Pr(a, b|0, 1)$ are shown in table I.

3. *Case 3: (x=1, y=0)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. For the state of the form $|0\rangle|\phi_0\rangle$, Bob obtains $|+\rangle$ with probability $\frac{1}{2}$ and $|-\rangle$ with probability $\frac{1}{2}$. In both the cases the second particle will be $|\phi_0\rangle$. For the state in form $|1\rangle|\phi_1\rangle$, Bob obtains $|+\rangle$ with probability $\frac{1}{2}$ and $|-\rangle$ with probability $\frac{1}{2}$. And for both the cases the second particle will be $|\phi_1\rangle$. Thus, the conditional probability $p(0, 0|1, 0) = p(M = |0'\rangle||+\rangle, |\phi_0\rangle) + p(M = |0'\rangle||+\rangle, |\phi_1\rangle)$, $p(0, 1|1, 0) = p(M = |1'\rangle||+\rangle, |\phi_0\rangle) + \Pr(M = |1'\rangle||+\rangle, |\phi_1\rangle)$ and $p(0, 2|1, 0) = p(M = |2'\rangle||+\rangle, |\phi_0\rangle) + \Pr(M = |2'\rangle||+\rangle, |\phi_1\rangle)$, where M is the measurement result of the second particle. Similarly, the conditional probability $p(1, 0|1, 0) = p(M = |0'\rangle||-\rangle, |\phi_0\rangle) + p(M = |0'\rangle||-\rangle, |\phi_1\rangle)$, $p(1, 1|1, 0) = p(M = |1'\rangle||-\rangle, |\phi_0\rangle) + p(M = |1'\rangle||-\rangle, |\phi_1\rangle)$ and $p(1, 2|1, 0) = p(M = |2'\rangle||-\rangle, |\phi_0\rangle) + p(M = |2'\rangle||-\rangle, |\phi_1\rangle)$. We accumulate all the conditional probabilities $p(a, b|1, 0)$ in table I.

4. *Case 4: (x=1, y=1)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 3. Conditional probabilities $p(a, b|1, 1)$ are shown in table I.

B. Conditional probability for entangled state with sub-systems lying in same subspace

Let us assume that Charlie supplies n entangled pairs of the form $\frac{1}{\sqrt{2}}(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle)$, where $|\phi_0\rangle = \cos\frac{\theta}{2}|i\rangle + \sin\frac{\theta}{2}|i+1\rangle$ and $|\phi_1\rangle = \cos\frac{\theta}{2}|i\rangle - \sin\frac{\theta}{2}|i+1\rangle$. Now, we analyze the case by case scenario below.

1. *Case 1: (x=0, y=0)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. For the first particle he obtains $|0\rangle$ with probability $\frac{1}{2}$ and in this case the second particle collapses to $|\phi_0\rangle$. When Bob measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, he obtains $|0'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8} + \sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$, $|1'\rangle$ with probability $\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8} - \sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$ and never gets $|2'\rangle$. Similarly, when Bob measures his first particle as $|1\rangle$, the second particle collapses to $|\phi_1\rangle$. In this case the probabilities of getting $|0'\rangle$, $|1'\rangle$ and $|2'\rangle$ are given in table II.

2. *Case 2: (x=0, y=1)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 1. Conditional probabilities $p(a, b|0, 1)$ are shown in table II.

3. *Case 3: (x=1, y=0)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. When he measures the first particle in $\{|+\rangle, |-\rangle\}$ basis he gets $|+\rangle$ with probability $\cos^2\frac{\theta}{2}$ and $|-\rangle$ with probability $\sin^2\frac{\theta}{2}$. In the first case, the second particle

TABLE I: Conditional probability of (a, b) given (x, y) for separable states

(x,y)	(a,b)	$\Pr((a,b) \mid (x,y))$
(0, 0)	(0, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}+\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(0, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}-\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}-\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(1, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}+\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(1, 2)	0
(0, 1)	(0, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{3\pi}{8}+\sin\frac{\theta}{2}\sin\frac{3\pi}{8})^2$
	(0, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{3\pi}{8}-\sin\frac{\theta}{2}\cos\frac{3\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{3\pi}{8}-\sin\frac{\theta}{2}\sin\frac{3\pi}{8})^2$
	(1, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{3\pi}{8}+\sin\frac{\theta}{2}\cos\frac{3\pi}{8})^2$
	(1, 2)	0
(1, 0)	(0, 0)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\cos^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\cos^2\frac{\pi}{8})$
	(0, 1)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\sin^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\cos^2\frac{\pi}{8})$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\cos^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\sin^2\frac{\pi}{8})$
	(1, 1)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\sin^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\cos^2\frac{\pi}{8})$
	(1, 2)	0
(1, 1)	(0, 0)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\sin^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\cos^2\frac{\pi}{8})$
	(0, 1)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\cos^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\sin^2\frac{\pi}{8})$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\sin^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\cos^2\frac{\pi}{8})$
	(1, 1)	$\frac{1}{2}(\cos^2\frac{\theta}{2}\cos^2\frac{\pi}{8}+\sin^2\frac{\theta}{2}\sin^2\frac{\pi}{8})$
	(1, 2)	0

collapses to $|0\rangle$. And in the second case, the second particle collapses to $|1\rangle$. When Bob measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, $|0'\rangle$ is obtained with probability $\cos^2\frac{\pi}{8}$, $|1'\rangle$ is obtained with probability $\sin^2\frac{\pi}{8}$and Bob never obtains $|2'\rangle$. Similarly, for $|1\rangle$, Bob obtains $|0'\rangle$ with probability $\sin^2\frac{\theta}{2}$, $|1'\rangle$ with probability $\cos^2\frac{\theta}{2}$ and never gets $|2'\rangle$. Conditional probabilities $p(a, b|1, 0)$ are shown in table II.

4. *Case 4: (x=1, y=1)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 3. Conditional probabilities $p(a, b|1, 1)$ are shown in table II.

C. Conditional probability for entangled state with sub-systems lying in different subspace

Let us assume that Charlie supplies n entangled pairs of the form $\frac{1}{\sqrt{2}}(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle)$, where $|\phi_0\rangle = \cos\theta|i+1\rangle + \sin\theta|i+2\rangle$ and $|\phi_1\rangle = \cos\theta|i\rangle - \sin\theta|i+2\rangle$. Now, we analyze the case by case scenario below.

1. *Case 1: (x=0, y=0)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. For the first particle he obtains $|0\rangle$ with probability $\frac{1}{2}$ and in this case the second particle collapses to $|\phi_0\rangle$. When Bob measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis, he obtains $|0'\rangle$ with probability $\cos^2\theta\sin^2\frac{\pi}{8}$, $|1'\rangle$ with probability $\cos^2\theta\cos^2\frac{\pi}{8}$ and $|2'\rangle$ with probability $\sin^2\theta$. Similarly, when Bob measures his first particle as $|1\rangle$, the second particle collapses to $|\phi_1\rangle$. In this case the probabilities of getting $|0'\rangle$, $|1'\rangle$ and $|2'\rangle$ are given in table III.

2. *Case 2: (x=0, y=1)*

In this case Bob measures his first particle in $\{|0\rangle, |1\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 1. Conditional probabilities $p(a, b|0, 1)$ are shown in table III.

3. *Case 3: (x=1, y=0)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. When he measures the first particle in $\{|+\rangle, |-\rangle\}$ basis he gets $|+\rangle$ with probability $\frac{1}{2}.\cos^2\theta$ and $|-\rangle$ with probability $\frac{1}{2}(1+\sin^2\theta)$. In the first case, the second particle collapses to $\frac{1}{\sqrt{2}}(|i\rangle+|i+1\rangle)$. And in the second case,

TABLE II: Conditional probability of (a, b) given (x, y) for entangled states with sub-systems in same space

(x, y)	(a, b)	$\Pr((a, b) \mid (x, y))$
(0, 0)	(0, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}+\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(0, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}-\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}-\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(1, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}+\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(1, 2)	0
(0, 1)	(0, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}+\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(0, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}-\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{2}(\cos\frac{\theta}{2}\sin\frac{\pi}{8}-\sin\frac{\theta}{2}\cos\frac{\pi}{8})^2$
	(1, 1)	$\frac{1}{2}(\cos\frac{\theta}{2}\cos\frac{\pi}{8}+\sin\frac{\theta}{2}\sin\frac{\pi}{8})^2$
	(1, 2)	0
(1, 0)	(0, 0)	$\cos^2\frac{\pi}{8}\cos^2\frac{\theta}{2}$
	(0, 1)	$\sin^2\frac{\pi}{8}\cos^2\frac{\theta}{2}$
	(0, 2)	0
	(1, 0)	$\sin^2\frac{\pi}{8}\sin^2\frac{\theta}{2}$
	(1, 1)	$\cos^2\frac{\pi}{8}\sin^2\frac{\theta}{2}$
	(1, 2)	0
(1, 1)	(0, 0)	$\sin^2\frac{\pi}{8}\cos^2\frac{\theta}{2}$
	(0, 1)	$\cos^2\frac{\pi}{8}\cos^2\frac{\theta}{2}$
	(0, 2)	0
	(1, 0)	$\cos^2\frac{\pi}{8}\sin^2\frac{\theta}{2}$
	(1, 1)	$\sin^2\frac{\pi}{8}\sin^2\frac{\theta}{2}$
	(1, 2)	0

the second particle collapses to $\frac{1}{\sqrt{2(1+\sin^2\theta)}}(-\cos\theta\,|i\rangle + \cos\theta\,|i+1\rangle + 2\sin\theta\,|i+2\rangle)$. Now, Bob measures the second particle in $\{|0'\rangle, |1'\rangle, |2'\rangle\}$ basis. In the first case, he gets $|0'\rangle$ with probability $\frac{1}{2}\cos^2\theta(\cos\frac{\pi}{8}+\sin\frac{\pi}{8})^2$, $|1'\rangle$ with probability $\frac{1}{2}\cos^2\theta(\sin\frac{\pi}{8}-\cos\frac{\pi}{8})^2$ and never gets $|2'\rangle$. For the second case, Bob obtains $|0'\rangle$ with probability $\frac{1}{2}\cos^2\theta(\sin\frac{\pi}{8}-\cos\frac{\pi}{8})^2$, $|1'\rangle$ with probability $\frac{1}{2}\cos^2\theta(\cos\frac{\pi}{8}+\sin\frac{\pi}{8})^2$ and $|2'\rangle$ with probability $\sin^2\theta$. All these conditional probabilities $p(a, b|1, 0)$ are shown in table III.

4. *Case 4: (x=1, y=1)*

In this case Bob measures his first particle in $\{|+\rangle, |-\rangle\}$ basis and second particle in $\{|0''\rangle, |1''\rangle, |2''\rangle\}$ basis. This case is similar to case 3. Conditional probabilities $p(a, b|1, 1)$ are shown in table III.

TABLE III: Conditional probability of (a, b) given (x, y) for entangled states with sub-systems in different subspaces

(x, y)	(a, b)	$\Pr((a, b) \mid (x, y))$
(0, 0)	(0, 0)	$\frac{1}{2}\cos^2\theta \sin^2\frac{\pi}{8}$
	(0, 1)	$\frac{1}{2}\cos^2\theta \cos^2\frac{\pi}{8}$
	(0, 2)	$\frac{1}{2}\sin^2\theta$
	(1, 0)	$\frac{1}{2}\cos^2\theta \cos^2\frac{\pi}{8}$
	(1, 1)	$\frac{1}{2}\cos^2\theta \sin^2\frac{\pi}{8}$
	(1, 2)	$\frac{1}{2}\sin^2\theta$
(0, 1)	(0, 0)	$\frac{1}{2}\cos^2\theta \cos^2\frac{\pi}{8}$
	(0, 1)	$\frac{1}{2}\cos^2\theta \sin^2\frac{\pi}{8}$
	(0, 2)	$\frac{1}{2}\sin^2\theta$
	(1, 0)	$\frac{1}{2}\cos^2\theta \sin^2\frac{\pi}{8}$
	(1, 1)	$\frac{1}{2}\cos^2\theta \cos^2\frac{\pi}{8}$
	(1, 2)	$\frac{1}{2}\sin^2\theta$
(1, 0)	(0, 0)	$\frac{1}{4}\cos^2\theta(\cos\frac{\pi}{8} + \sin\frac{\pi}{8})^2$
	(0, 1)	$\frac{1}{4}\cos^2\theta(\sin\frac{\pi}{8} - \cos\frac{\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{4}\cos^2\theta(\sin\frac{\pi}{8} - \cos\frac{\pi}{8})^2$
	(1, 1)	$\frac{1}{4}\cos^2\theta(\cos\frac{\pi}{8} + \sin\frac{\pi}{8})^2$
	(1, 2)	$\sin^2\theta$
(1, 1)	(0, 0)	$\frac{1}{4}\cos^2\theta(\cos\frac{\pi}{8} + \sin\frac{\pi}{8})^2$
	(0, 1)	$\frac{1}{4}\cos^2\theta(\sin\frac{\pi}{8} - \cos\frac{\pi}{8})^2$
	(0, 2)	0
	(1, 0)	$\frac{1}{4}\cos^2\theta(\sin\frac{\pi}{8} - \cos\frac{\pi}{8})^2$
	(1, 1)	$\frac{1}{4}\cos^2\theta(\cos\frac{\pi}{8} + \sin\frac{\pi}{8})^2$
	(1, 2)	$\sin^2\theta$

Construction and Building Materials 175 (2018) 682–690

Contents lists available at ScienceDirect

Construction and Building Materials

journal homepage: www.elsevier.com/locate/conbuildmat

Protective polyurea coatings for enhanced blast survivability of concrete

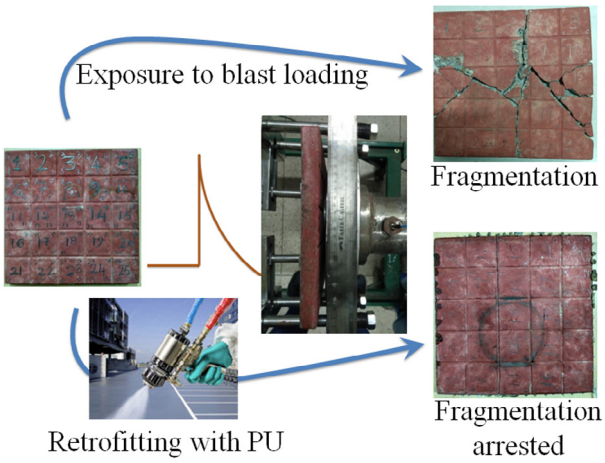
N. Iqbal [a,b], P.K. Sharma [a], D. Kumar [b], P.K. Roy [a,*]

[a] Centre for Fire, Explosive and Environment Safety, DRDO, Timarpur, Delhi 110054, India
[b] Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi 110042, India

HIGHLIGHTS

- A chain-extended polyurea has been coated on concrete substrate.
- The efficacy of polyurea towards shock loadinghas been demonstrated.
- The effect of coating thicknesstowards blast survivability has been studied.
- Detailed mechanism for explaining the improved blast resistance has been discussed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 February 2018
Received in revised form 6 April 2018
Accepted 24 April 2018

Keywords:
Polyurea
Concrete
Dynamic glass transition
Blast mitigation
Shock tube

ABSTRACT

Polyurea is a synthetic high-strength elastomeric coating that can be sprayed over existing structures for protection against weathering effects. In view of the fast kinetics associated with the isocyanate-amine reaction, practical processing of polyurea is performed using spray coating technique. In this paper, we report the effect of varying the type and amount of chain extender on the mechanical properties of spray coated polyurea. An optimal ratio of aromatic: aliphatic chain extender was found to result in optimal H-bonding, which in turn reflected in terms of improved mechanical properties of the coating. In this paper, the efficacy of polyurea towards improving the performance of the substrate towards shock and ballistic loading is demonstrated. Polyurea coated concrete tiles were subjected to controlled blast loadings, and an attempt was made to understand the mechanism responsible for the improved blast survivability. Unreinforced concrete tiles underwent extensive fragmentation at peak pressures <50 psi, while polyurea coated tiles could withstand much higher peak pressures. The extent of mitigation increased with the coating thickness and composites with 6 mm polyurea could withstand substantially higher pressures. Upon further loading, polyurea-concrete debonding was evidenced, however the membrane arrested the fragments formed. Dynamic mechanical studies revealed that polyurea remains as an elastomer under the frequency range associated with the shock tube testing and behaves as a catcher system for the fragments formed due to blast loading.

© 2018 Elsevier Ltd. All rights reserved.

* Corresponding author.
E-mail address: pkroy@cfees.drdo.in (P.K. Roy).

https://doi.org/10.1016/j.conbuildmat.2018.04.204
0950-0618/© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Polyurea is a polymerization product resulting from the reaction between diisocyanates (—NCO) and diamines ($—NH_2$). In view of the fast kinetics associated with this reaction, practical processing of polyurea is only possible through a reactive spray coating technique, where the reactants are stored separately and directed to react in the mixing-module, just prior to being pushed out under pressurized conditions through an orifice [1]. Spray coating yields a continuous membrane devoid of seams and joints, which improves the abrasion, corrosion and chemical resistance of structures. In view of the zero Volatile Organic Content (VOC) content of polyurea, the Environmental Protection Agency (EPA) has also approved their use as a coating over concrete.

What renders polyurea particularly interesting is the possibility to tune its mechanical properties by tinkering with the reaction chemistry involving the isocyanate, amine and the chain extender. The properties of the polyurea coating can further be enhanced by tuning the processing conditions, and most commercial formulations are sprayed at elevated temperatures ($\geq$70 °C) and pressures $\sim$2000 psi or higher.

In addition to enhance the structural performance of concrete [2–4], coating with polyurea has been found to lead to improvement in the blast survivability as well. Most of the existing masonry structures were originally designed to withstand static loads and in view of the increasing terrorist threats, it is highly desirable for all vulnerable buildings to possess an additional capability to withstand dynamic loads. In this context, preliminary studies were performed by the Air Force Research Laboratory, where polyurea coating was used towards strengthening of masonry and light weight steel structures [5]. Coating polyurea on the interior face of walls was found to be more beneficial, in comparison to the blast facing side and both side application was found to be economically prohibitive [6]. Polyurea reinforced masonry structures, upon being subjected to transient dynamic loads, were found to undergo large deflections thereby endowing the much-needed ductility during flexural failure of the masonry walls [7]. Polyurea coatings have also been reported to improve the survivability of substrates like steel plates [8–10] and fiber-reinforced face-sheets [11–15].

In view of the commercial aspects involved, the details of polyurea formulations are unavailable in open literature. Polyurea formulations are usually available as dual component system: part 'A' being the isocyanate and part 'B' being the amine resin blend. In addition, the latter component also contain chain extenders, however, there are no guidelines to delineate their optimal amounts, which result in superior performance of the material. In this paper, we establish the effect of nature and amount of chain extender on the material properties of polyurea formulation. Concrete tiles were subsequently coated with the developed polyurea formulation with an aim to establish the effect of increasing coating thickness towards improving its blast survivability.

Another aim of the present work is to establish the underlying contributory mechanism behind the role of polyurea towards improving the blast resistance of the composite. Several mechanisms have been reported [14,16–19], including a dynamic "Rubber-glass transition" where a vitrification process reportedly transpires in the polyurea coating upon being subjected to blast loads [16,20]. However, we hypothesize that the occurrence of such a dynamic transition is unlikely, as exceptionally high frequencies would be required for its incidence: conditions not achievable under blast loadings. Such conditions can be simulated by systematic dynamic mechanical studies over a wide range of frequencies. In the present study, time–temperature superposition (TTS) is used to establish the mechanical properties of polyurea using the dynamic mechanical analysis data, since direct measurement over such a broad frequency range is not possible. Of particular interest is to identify the frequency associated with the dynamic loading process, which can initiate the dynamic 'glass transition' process in polyurea coating.

2. Experimental

2.1. Materials

Concrete tiles conforming to IS 8112-1989 (M35 Grade, 28 day compressive strength 35 MPa) were used as the underlying substrate for polyurea spray coating. The details of the concrete used for the present work are presented in the supplementary section (Section S1). Commercial isocyanate prepolymer, Suprasec 2054 (Huntsman), NCO = 15.5% was used without any further purification. Poly(propylene oxide) (PPO) based polyether amines JEFFAMINE®D-230 and JEFFAMINE®D-2000 were procured from Huntsman: the designation number referring to the approximate molecular weight of each species, with D referring to the functionality (D: difunctional). An aromatic chain extender, diethyltoluenediamine (DETDA 80, Lonzacure™) was used without any further purification. The same has been reported to be a mixture of 3,5-diethyltoluene-2,4-diamine (2,4-DETDA) (75–81%), 3,5-diethyltoluene-2,6-diamine (2,6-DETDA) (18–20%) the remaining being dialkylated m-phenylenediamine (0.5–3%). Coatosil MP 200 (Momentive) was included in the formulation to improve the adhesion of the coating with the concrete tile. The structure of the reagents used for preparation of polyurea coating are presented in Table 1.

2.2. Polyurea processing

Suprasec 2054, duly diluted with propylene carbonate, was used as isocyanate, and a formulated mixture containing requisite amounts of D-230, D-2000 and DETDA was used as the amine blend for polyurea preparation. The amount of each of the component was calculated as follows:

$$m_{iso} \cdot n_{eq.iso} = [m_{amine} \cdot n_{eq.amine}] i_{NCO} \quad (1)$$

where, m_{iso} and m_{smine} refer to the mass of isocyanate and amine respectively and $n_{eq,iso}$ and $n_{eq,amine}$ are the equivalent number of isocyanate and amine respectively. The isocyanate index (i_{NCO}) refers to the desired NCO index, which was maintained at 1.1for all formulations. The amount of amine (parts by volume) used for preparing different polyurea formulations are presented in Table 2, with the volume percentage of Suprasec 2054 and propylene carbonate being 90.8 and 9.2 respectively for each formulation. The amount of the hard segment content (HS) was determined as per established procedure (Eq. (2)) [21].

$$HS = \frac{m_{iso} + m_{chain\ extender}}{m_{iso} + m_{chain\ extender} + m_{amine}} \quad (2)$$

It is to be noted that the m_{iso} used for the calculation of hard segment refers to the mass of isocyanate prepolymer which also includes the contribution of the long polyether chain. Detailed calculations involving the individual contributions of amines and the chain extenders to the total urea present in the sample is presented in the supplementary section, and have been tabulated and depicted pictorially in Table S1 and Fig. S1 respectively. The samples have been designated as PU followed by the (%) contribution of aromatic chain extender in the polymer.

Table 1
Structure and physico-chemical properties of reactants.

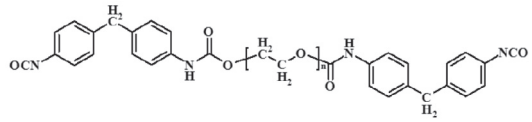
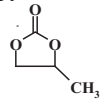
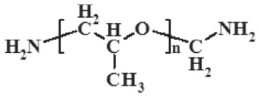
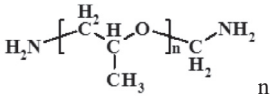
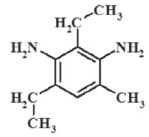
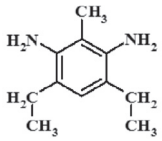
Structure	Density (g/cm^3)	n_{eq}	Viscosity (30 °C, mPas.S)
Part A: Isocyanate diluted with propylene carbonate			
Suprasec 2054 	1.09	3.57	878.3
Propylene carbonate 	1.2	–	2.2
Part B: Amine blend			
Jeffamine D 2000  n ~ 33	0.98	1	182.9
Jeffamine D 230  n ~ 3	0.93	8.1–8.7	7.46
3,5-diethytoluene-2,4-diamine 	1.022	11.24	106.8
3,5-diethytoluene-2,6-diamine 			

Table 2
Details of polyurea formulations prepared.

Sample designation	Amine (part per unit volume)			Total Hard segment (%)	Contribution of aromatic chain extender (%)
	Chain Extender		D2000		
	D 230	DETDA			
PUAr-76	–	21.22	78.78	60.86	76
PUAr-63	5.48	17.52	77.0	61.65	63
PUAr-36	16.41	11.05	72.54	63.72	36
PUAr-10	27.35	2.68	69.97	64.84	10
PUAr-0	31.30	–	68.70	65.43	–

2.3. Coating on concrete tiles

Coating performance is strongly influenced by its ability to adhere properly to the substrate material, which in turn is dependent on the surface preparation. Proper surface preparation allowed creation of suitable profile for strong mechanical bond to exist between polyurea and the underlying concrete substrate. Concrete tiles (25 cm × 25 cm × 2 cm) were cleaned to remove excessive cement slurry, laitance, loose and crumbly particles, joints and oily substances. The substrate was primed with epoxy following which the polyurea formulation was sprayed using a spray-coating machine (E 10 HP equipped with a fusion air-purge spray gun AW2222, Graco). In all the runs, block and hose heaters were maintained at a constant temperature of 75 °C and the spraying operation was performed at 2000 ± 100 psi.

2.4. Characterisation techniques

Mechanical properties were measured in tensile mode using Universal Testing Machine (International Equipments) as per ASTM D 412C standard (50 mm/min). Tear strength was determined as per ASTM D 624C (loading rate 500 mm/min). At least five specimens were tested and the average results along with the standard deviation values have been reported.

Dynamic mechanical analysis was performed using Netzsch Instrument Dynamic Mechanical Analyzer (Artemis DMA 242E). Dual-cantilever bending mode was employed for testing, which was analysed using the corresponding data analysis software. The temperature sweep is conducted at four different frequencies (1 Hz, 5 Hz, 10 Hz and 20 Hz) simultaneously. The temperature was ramped at a rate of 2 °C/min with the deformation occurring

at the given frequencies from −80 °C to 150 °C which was expected to cover the glass transition temperature (T_g), and most of the transition zone between the glassy state and the rubbery state. The nominal sample dimensions for DMA test is 50 mm × 8 mm × 3 mm (l × b × h). A dual cantilever setup was used for testing, which permitted controlled movement of the sample during testing, while the cantilever fixtures clamp the sample at both ends as well as the centre.

Thermal behavior of samples was investigated using Perkin Elmer Diamond STG-DTA in N_2 atmosphere (flow rate = 20 mL/min) over the temperature range of 50–800 °C. A heating rate of 10 °C/min and sample mass of 5.0 ± 0.5 mg was used for each experiment. Differential Scanning Calorimeter (TA instruments Q 20) was used for calorimetric studies. DSC scans were performed under flowing N_2 (50 mL/min) to minimize oxidative degradation of the sample. Requisite amount of sample (10 ± 2 mg) was sealed in aluminum pans, and then heated from 50 to 200 °C at 10 °C/min. FTIR spectra of samples were recorded in the wavelength range 4000–600 cm^{-1} using Fourier Transform Infrared (FTIR) spectroscopy on a Thermo Fisher FTIR (NICOLET 8700) with an attenuated total reflectance (ATR) crystal accessory. The adhesive strength of the Polyurea coated concreted samples was measured according to ASTM D 4541using an adhesion tester (Elcometer 510T) with a 20 mm diameter dolly. An epoxy-based adhesive (Araldite) was used to fix the dolly to the polyurea coating.

2.5. Shock loading system

A shock tube facility at the Centre for Fire, Explosive and Environment Safety, was employed to subject the coated concrete tiles to dynamic loadings. The shock tube consists of a long rigid cylinder, divided into a high-pressure driver section and a low pressure driven section, separated by a Mylar diaphragm; which is made to rupture at pre-determined pressure differences. The rapid release of gas results in the creation of a shock wave, which travels down the tube to impart blast dynamic loadings on the specimen. The shock tube facility has an overall length of 8.7 m, which is divided into a 1.7 m driver section and a 7 m driven section. A schematic is presented in the supplementary section (Fig. S2). Pressurized air is used as the driver gas, the driven gas being ambient air. Pressure sensor (113B21, PCB make), placed at a distance of 11.3 cm from the sample was used to capture the pressure profile during the testing. Strain gauge (HBM: K-LY4-1-11-120-3-2) was affixed on the blast–opposing face of the concrete tile to measure the deflection in the sample.

2.5.1. Concrete tile geometry and boundary conditions

Concrete tiles (25 cm × 25 cm × 2 cm) were coated with polyurea of varying thickness (1–6 mm) and subjected to controlled blast loadings. All tiles were simply supported over a span of 10 cm along any two edges, with the remaining edges free. A circular region of 10 cm diameter was effectively subjected to blast loading, as pictorially shown in Fig. 1.

3. Results and discussion

The effect of chain extension process on the properties of polyurea has been systematically studied to develop a spray processable polyurea formulation with optimal properties. The mechanical response of polyurea coated concrete tiles subjected to controlled blast loadings was studied to establish the effect of increasing thickness on the blast survivability and to gain an insight into the underlying mechanism responsible for the improved blast resistance.

3.1. Mechanical properties of polyurea

For the development of sprayable formulations, the isocyanate part was duly diluted with a reactive diluent, namely propylene carbonate [22]. The rheological profile of both the parts, i.e. isocyanate and formulated amine is presented in the supplementary section (Fig. S3). It is clear that both sides exhibited similar viscosities over the temperature range of spray processing (T >70 °C), which permits maintenance of stoichiometric conditions (1:1) in the mixing module.

Uniform polyurea films (thickness ~3 mm) were prepared on Teflon sheets, which were removed for evaluation of mechanical properties. The effect of increasing amount of aromatic chain extender (DETDA) and aliphatic chain extender (D230) on characteristic parameters, namely 'gel-time' and 'tack free' time are presented in the supplementary section (Table S2). Formulations containing larger amount of aliphatic chain extender (D230) required relatively longer time periods for the flow to stop. Interestingly, although the solidification took place in a matter of seconds, the mechanical properties improved significantly with time: a phenomenon which can be attributed to the release of internal stresses built up within the polymer during processing conditions. The time dependent emergence of mechanical properties for a representative formulation is presented in Fig. 2. It can be seen that the properties evolved significantly over time and reached a plateau after ~15 days. This improvement can be attributed to several processes simultaneously occurring within the matrix; the foremost being the relaxation of internal stresses which tend to buildup in the polymer during the spray coating process [23,24]. In addition, significant changes occur in the internal morphology of segmented polymers, which in turn is a result of H-bond rearrangement [25,26]. The evolution of mechanical properties for the remaining formulations is presented in the supplementary section (Figs. S4 and S5).

Characteristic mechanical properties of fully cured polyurea coatings (post 15 days of curing) in terms of tensile strength, elongation and tear strength are presented in Fig. 3. The individual contributions of the long chain amine (D2000), aromatic (DETDA) and aliphatic chain extender (D230) are also included for ready reference.

It is to be noted that the interesting properties of polyurea are basically an indirect manifestation of its internal microstructure. The extensive H-bonding between the macromolecular chains leads to the segregation of the polymer into separate domains [27]: "hard" realms formed by hydrogen bonded polar urea linkages (—NH—CO—NH—) and "soft" domains comprising of long aliphatic polyether chains [28,29]. In the absence of chain extender, the urea linkages in polyurea are placed relatively farther from each other, which hinder the formation of H-bonds. Chain extension increases the vicinity of the urea groups (Fig. S6), which favor their further alliance into hard domains.

It is clear that an optimal ratio of aromatic: aliphatic chain extender directs the growing polyurea macromolecule in such a way so as to result in optimal H-bonding. Polyurea formulation with 61% hard segment (aromatic contribution to urea linkages = 63%) exhibited best properties (Tensile strength 13.7 ± 0.7 MPa, elongation 229 ± 10%, Tear strength 87.78 ± 5.8 N/mm). Hence, the formulation was sprayed over concrete tiles, which in turn were subjected to adhesion tests followed by shock tube experiments. It is to be noted that for any coating to perform its role as an effective protective layer, a high level of adhesion is essential. The adhesion strength could not be determined experimentally, due to the cohesive failure of the epoxy adhesive between polyurea film and the pull-off adhesive test dolly at ~13 MPa. The concrete-polyurea debonding was not observed under the experimental conditions employed for adhesion testing.

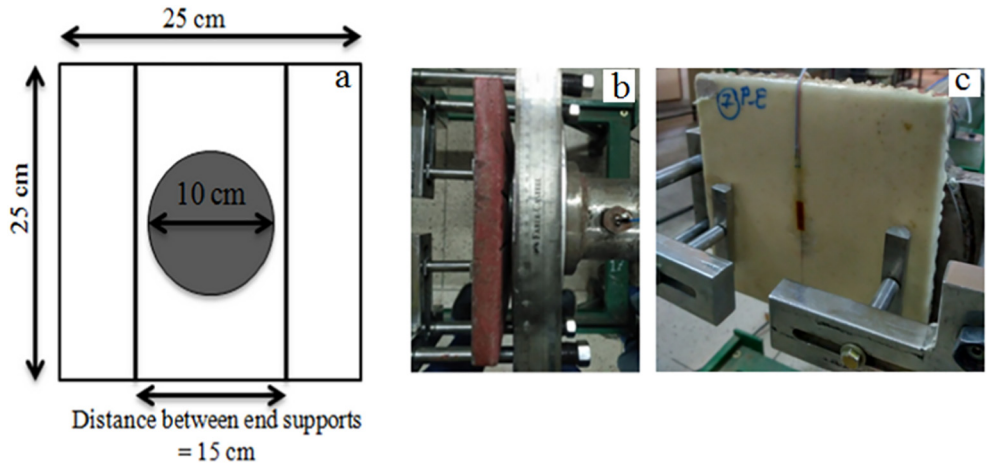


Fig. 1. a) Sketch of the specimen depicting the loading and boundary areas. (b) Actual specimen placement c) Front view.

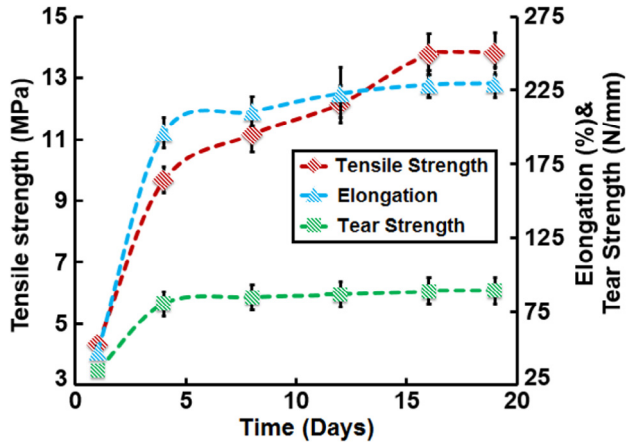


Fig. 2. Evolution of tensile properties as a function of time (PU-Ar 63) Tensile strength – Primary axis, Elongation and tear strength – secondary axis.

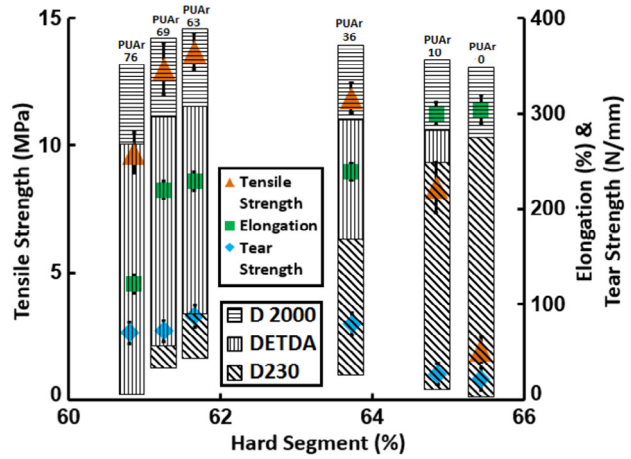


Fig. 3. Variation in tensile strength, elongation and tear strength of polyurea due to chain extension. The relative contribution of aromatic and aliphatic hard segment is also included in the figure.

3.2. Blast performance of polyurea coated concrete tiles

The blast response of polyurea-coated tiles was evaluated by subjecting them to controlled blast loadings on a shock tube followed by macroscopic visual examinations and strain measurements. A shock tube facility was employed to subject coated concrete tiles to blast loading conditions. The pressure profile obtained at the sensors located at 11 cm from the sample is recorded and a representative profile is presented in the supplementary section (Fig. S7). A set of control experiments on uncoated tiles were also performed for direct evaluation of polyurea retrofitting effectiveness. In the present context, "effectiveness" is defined as the ability of polyurea to prevent failure and to arrest the fragments formed from being propelled, if any. Table 3 summarizes the values for reflected overpressure, impulse, strain along with the summary of observations, and post-blast visuals.

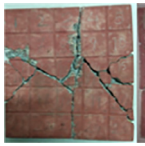
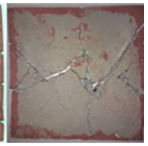
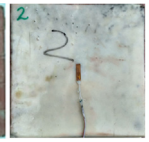
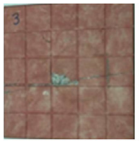
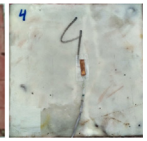
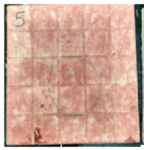
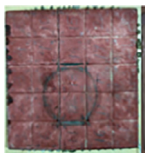
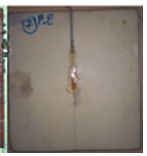
Neat concrete tiles were found to fail at a reflected shock pressure of ~50 psi, and the fragments formed were propelled at high velocities. It is to be noted that most casualties sustained by inhabitants during external explosions are not caused by heat/overpressure. Instead, it is the disintegration and fragmentation of walls followed by propulsion of non-secured objects, which lead to blunt trauma and penetration injuries. Retrofitting with polyurea coating can significantly reduce such damages, however the coating thickness may vary depending on the type of structure.

Coating the concrete tile with polyurea led to arrest of fragments and the extent of mitigation was found to be proportional to the coating thickness. Coating of tile with polyurea (thickness ~1 mm) was found to be effective in holding back the fragments formed, when subjected to higher blast loads. The retrofitted tile (6 mm thickness) finally underwent fragmentation at ~90 psi, with concrete-polyurea debonding and failure of strain gauge. Earlier studies on polyurea coated walls (~6 mm thickness) were found to resist peak pressures greater than ~60 psi for one-way action walls, compared to a capacity of ~5 psi for unreinforced masonry walls [7].

3.3. Mechanism behind improved blast resistance

The pressure-time profile associated with the shock tube experiments is presented in the supplementary section (Table S3). Frequencies, associated with the blast loading experiments were estimated as the inverse of rise-time (t_{rise}) in pressure–time profile. Under the experimental conditions, t_{rise} was found to vary between 2 and 2.5 ms, i.e. blast frequency (ν_{blast}) of ~400–500 Hz. In fact, in most of the shock tube tests, the rise-time is of the same order [15,30]. To gain an insight into the underlying mechanism behind the improved blast survivability of coated tiles, the mechanical

Table 3
Response of polyurea coated concrete tiles.

PU thickness (mm)	Peak pressure/ impulse	Observations	Visual inspection (post blast)	
			Blast face	Opposite face
Control (Nil)	50 psi/404 psi.ms	Extensive fragmentation was evidenced, with the fragments being propelled at high velocities (expected to cause harm to occupants on the other side)		
1	70 psi/406 psi.ms	Strain: 165 με Cracks formed on the tile. Slight tearing of polyurea, but the polyurea membrane proved effective in holding the tile together. Concrete-polyurea debonding was not evidenced		
2	72 psi/435 psi.ms	Strain: Nil A single line crack formed on the tile. Polyurea tearing was not observed and the membrane was effective in holding the tile together. Tearing of the polyurea membrane was not evidenced		
3	74 psi/601 psi.ms	Same as above		
4	74 psi/610 psi.ms	Same as above		
5	82 psi/724 psi.ms	Strain: Nil The coated tile was intact with no visible cracks even on the blast facing side. Concrete-polyurea debonding was not observed. Tearing of the polyurea membrane was not evidenced		
6	86 psi/798 psi.ms	Strain: 234 με The coated tile was intact with no visible cracks even on the blast facing side. Concrete-polyurea debonding was not observed. Tearing of the polyurea membrane was not evidenced		

response of the fully cured polyurea (PU-Ar 63) under dynamic loading conditions was studied in detail.

The complex dynamic Young's modulus E*, as obtained from DMA studies, can in turn be represented as $E^* = E' + iE''$, where E′ refers to the storage modulus, representing the stiffness of the material, and E″ is the loss modulus, representing the ability of the material to dissipate mechanical energy. The former is indicative of the energy stored in the material during the cyclic loading process, which is subsequently released during unloading. The ratio of loss and storage moduli (tan δ) is representative of the ratio of the dissipated and stored mechanical energy in each loading cycle.

The variation of storage and loss modulus of PU Ar 63 at various frequencies as a function of temperature is presented in Fig. 4. The loss factor (tan δ) is also presented. A single relaxation peak is clearly evident in the profile, which is attributed to the segmental motions linked with the soft segments in the polymer. A sub-zero T_g (peak of the loss modulus) [31,32]was evidenced at −36 °C (@1Hz), which leads to an elastomeric response of polyurea under ambient conditions. It is interesting to observe the large width associated with the glass transition region (ΔT ~80 °C), which is a property rather unique to polyurea [33]. As the frequency of the dynamic loading is increased, both the loss and storage modulus trace shift towards higher temperature [31–34].

For most materials, increasing the temperature, and decreasing speed of testing (frequency) have the same effect on its mechanical properties, especially strength [31]. This relationship permits evaluation of mechanical properties over a wide frequency range by conducting tests at relatively narrow frequencies but at various temperatures. This allows the prediction of specific mechanical properties by shifting the response in the form of curve obtained at various temperatures. This method, commonly known as the Time–Temperature Superposition (TTS) was applied to arrive at the frequency domain master curves, spanning over 10^{-15} to 10^{20} Hz. The master curves hence obtained in terms of loss modulus variation at different reference temperatures (245 K–298 K) are presented in Fig. 5.

It is evident from the figure that a significant shift in the loss modulus peak (encompassing dynamic glass transition phenomenon) occurs upon increasing temperature. This can be attributed to the fact that segmental motions at higher temperatures mandate relatively higher frequencies to get arrested [16]. In fact, this 'vitrification' process has been considered to be one of the most contributory mechanism behind the ability of polyurea to withstand blast loadings [16,20].

Interestingly, our present studies indicate that under ambient temperatures, the dynamic glass transition process necessitate far high frequencies (~10^{13} Hz) than those accomplishable under shock tube conditions. Therefore, instead of undergoing a vitrification process, the polyurea retains its elastomeric nature and flexes to act as a "catcher system" for the fragments formed due to the fracture of the underlying concrete.

Polyurea helps in the dissipation and absorption of shock wave kinetic energy by undergoing irreversible densification, which in turn results from the hard domain ordering and crystallization. These micro structural changes and H-bond restructuring reportedly dissipate and absorb significant amount of shock-wave kinetic energy [18]. Theoretical studies have also brought out the providence of "shock wave capture and neutralization" mechanism in polyurea [14,18,29,35]. Shock waves usually travel as a single wave in homogeneous materials. However, due to the heterogeneous microstructure of polyurea, a two-wave structure is obtained—a leading shock front followed by a time variable complex pattern. It is this impedance mismatch over the length scales and nonlinearities arising from material inelasticity, which leads to a dual shock-wave pattern. The secondary trailing shock-wave (release

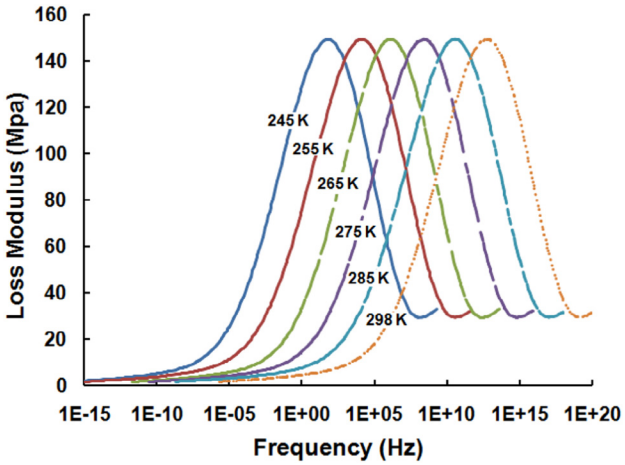


Fig. 5. Master curves representing the variation in loss modulus in PU-Ar 63.

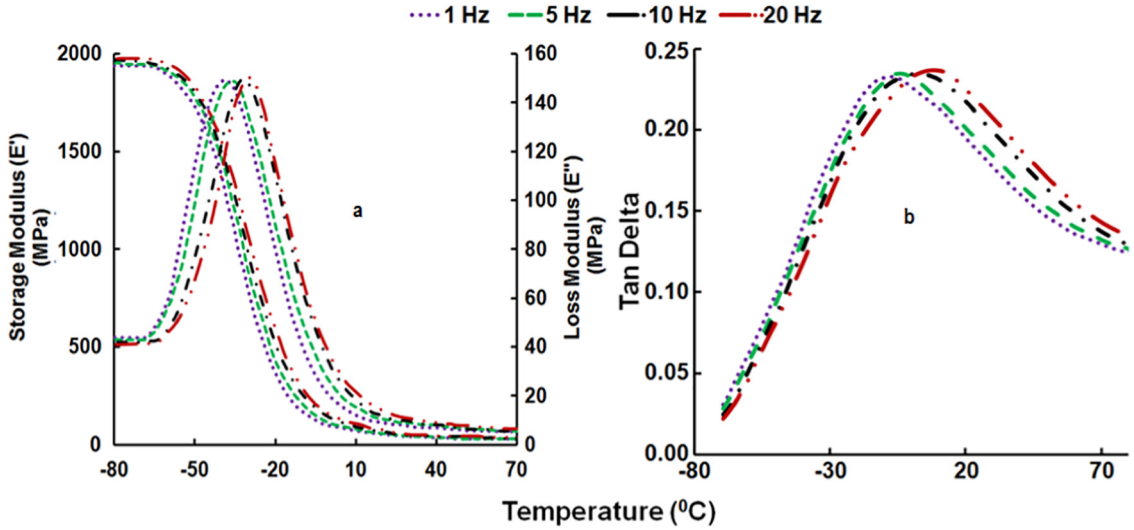


Fig. 4. Variation of (a) Storage and loss modulus (b) Tan delta with temperature at different frequencies.

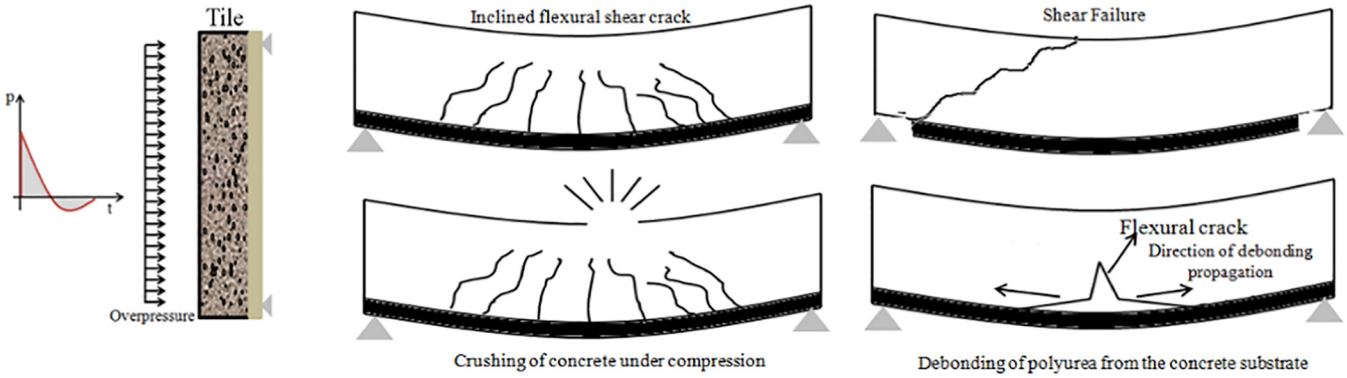


Fig. 6. Failure mechanisms in retrofitted plain cement concrete.

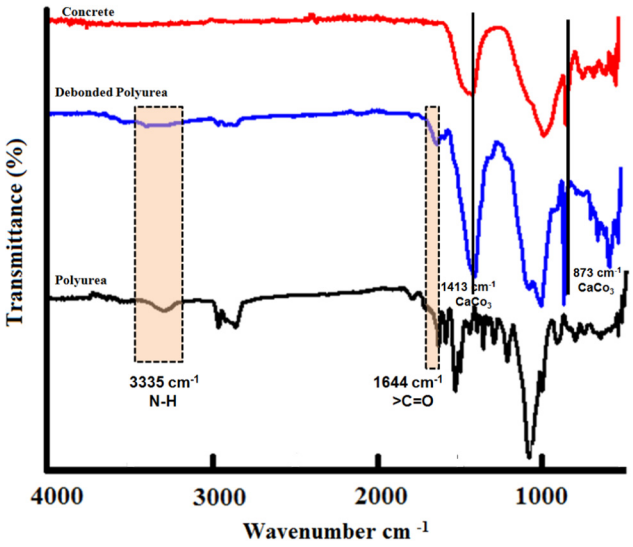


Fig. 7. FTIR spectra of polyurea, concrete and debonded polyurea.

wave), catches up and attenuates the leading shockwave, thereby leading to shock attenuation [36].

It is to be noted that the blast facing side of the concrete tile is exposed to compressive loadings, while the opposite face is subjected to tensile forces. Plain Cement Concrete (PCC) exhibits high compressive strengths (σ_c) with the tensile strength (σ_t) being only a small fraction ($\sigma_t = 10\%\ \sigma_c$) [37]. As the shock wave propagates, the structure is weakened and the stresses localize at the weak sections. Concrete is an essentially heterogeneous material consisting of aggregates and cement bonded together, with lots of inherent weak sites capable of acting as initiating sites for subsequent failure. The primary modes of failure in the retrofitted tile are presented in Fig. 6.

Due to high local tensile stresses, "inclined flexural shear cracks" propagate within the structure, resulting in shear failure. The flexural failure is accompanied with the crushing of concrete, which occurs primarily on the side facing the blast load. As the fragmented tile flexes (Fig. S8), polyurea is subjected to tensile forces, which leads to tearing of the coating. This was observed only when the coating thickness was less (~1 mm).

The overall efficiency of the polyurea-coated concrete system is limited by the adhesive strength between the polyurea and the concrete surface. For all practical purposes, the cohesive strength of concrete is much lower in comparison to the adhesive strength between polyurea and the substrate [38]. Inclined microcracks are initiated in the superficial layer of concrete in view of its lower tensile strength as compared to the adhesive (Fig. S8). These, however, cannot propagate more than few millimeters within the concrete specimen as the stresses decrease very rapidly with depth along polyurea-concrete interface. Concrete-polyurea debonding was evidenced at higher loadings, when the coated tile (6mm PU coating thickness) was subjected to 90 psi overpressure, with ~3 mm of substrate adhering to the polyurea membrane. A visual of the debonded polyurea is presented in the supplementary section (Fig. S9).

To understand the nature of the interface, FTIR studies were performed, and the spectra of concrete, polyurea and debonded surface of polyurea is presented in Fig. 7. The detailed assignment of the spectral bands is presented in the supplementary section Table S4. Polyurea exhibits characteristics absorption peaks at 1644 cm^{-1}due to urea carbonyl (ordered) [1] and absorption at 3335 cm^{-1} due to the NH vibration. In the spectra of concrete, characteristic absorption bands were observed at 797 and 777 cm^{-1} due to Si—O vibrations. Additional peaks at 1413 and 873 cm^{-1}were observed, which could be attributed to the presence of calcium carbonate C═O further confirming the presence of the adhered cementitious layer onto the polyurea coating [39].

4. Conclusions

The mechanical property of polyurea was tuned by introducing aliphatic and aromatic chain extenders in the amine co-reactant side of the formulation. The presence of chain extender increases the vicinity of urea linkages, which in turn favors the formation of H-bonds in the hard segments. An optimal ratio of the two forces the growing macromolecule to orient in a way so as to result in optimal H-bonding, which reflect in terms of improved mechanical properties. An experimental study towards evaluating the response of polyurea coated concrete tiles under controlled blast loading conditions in a shock tube was further performed. The flexible nature of polyurea permitted it to flex, thereby preventing the fractured concrete tiles from turning into secondary fragment hazards. Detailed dynamic mechanical studies clearly indicated that a minimum of 10^{13} Hz was required for the polyurea to undergo glass transition under ambient temperatures. Interestingly, the frequency associated with the blast loading conditions ($\sim 10^2$ Hz) was found to be insufficient to permit the vitrification of polyurea vide the much-cited "dynamic glass transition" phenomenon. Instead, shock-wave-induced hard domain ordering and crystallization, rearrangements in the H-bonds and shock wave capture and neutralization appear to be the primary modes of energy dissipation in polyurea during blast loading conditions.

Conflict of interest

We confirm that there is no conflict of interest.

Acknowledgements

Rajan is thanked for their help during operation of the shock tube facility. The authors thank DRDO for funding this work through ST/14-15/CFE-1311.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.conbuildmat.2018.04.204.

References

[1] Q. Zhou, L. Cao, Q. Li, Y. Yao, Z. Ouyang, Z. Su, X. Chen, Investigation of the curing process of spray polyurea elastomer by FTIR, DSC, and DMA, J. Appl. Polym. Sci. 125 (5) (2012) 3695–3701.
[2] H.A. Toutanji, H. Choi, D. Wong, J.A. Gilbert, D.J. Alldredge, Applying a polyurea coating to high-performance organic cementitious materials, Constr. Build. Mater. 38 (2013) 1170–1179.
[3] L. Ping, Xin Mao Li, W.B. Huang, Effect of dry-wet circulation and temperature change on properties of polyurea coatings, Adv. Mater. Res. 150–151 (2011) 1203–1208.
[4] S. Parniani, H. Toutanji, Monotonic and fatigue performance of RC beams strengthened with a polyurea coating system, Constr. Build. Mater. 101 (2015) 22–29.
[5] K.J. Knox, M.I. Hammons, T.T. Lewis, J.R. Porter, Polymer Materials for Structural Retrofit, Florida, 2000.
[6] J.S. Davidson, J.W. Fisher, M.I. Hammons, J.R. Porter, R.J. Dinan, Failure mechanisms of polymer-reinforced concrete masonry walls subjected to blast, J. Struct. Eng. 131 (8) (2005) 1194–1205.
[7] J. Davidson, J. Porter, R. Dinan, M. Hammons, J. Connell, Explosive testing of polymer retrofit masonry walls, J. Perform. Constr. Fac. 18 (2) (2004) 100–106.
[8] K. Ackland, C. Anderson, T.D. Ngo, Deformation of polyurea-coated steel plates under localised blast loading, Int. J. Impact Eng. 51 (2013) 13–22.
[9] K. Ackland, C. Anderson, N. St John, Polymeric Coatings for Enhanced Protection of Structures from the Explosive Effects of Blast, RNSA Security Technology Melbourne, Australia, 2007, pp. 90–96.
[10] L. Cai, A. Al-Ostaz, X. Li, C. Fowler, A.H.D. Cheng, H. Alkhateb, Protection of steel railcar tank containing liquid chlorine from high speed impact by using polyhedral oligomeric silsesquioxane-enhanced polyurea, Int. J. Impact Eng. 75 (Supplement C) (2015) 1–10.
[11] A. Grujicic, M. LaBerge, M. Grujicic, B. Pandurangan, J. Runt, J. Tarter, G. Dillon, Potential improvements in shock-mitigation efficacy of a polyurea-augmented advanced combat helmet, J. Mater. Eng. Perform. 21 (8) (2012) 1562–1579.
[12] M. Grujicic, W.C. Bell, B. Pandurangan, P.S. Glomski, Fluid/structure interaction computational investigation of blast-wave mitigation efficacy of the advanced combat helmet, J. Mater. Eng. Perform. 20 (6) (2011) 877–893.
[13] M. Grujicic, W.C. Bell, B. Pandurangan, T. He, Blast-wave impact-mitigation capability of polyurea when used as helmet suspension-pad material, Mater. Des. 31 (9) (2010) 4050–4065.
[14] M. Grujicic, R. Yavari, J.S. Snipes, S. Ramaswami, J. Runt, J. Tarter, G. Dillon, Molecular-level computational investigation of shock-wave mitigation capability of polyurea, J. Mater. Sci. 47 (23) (2012) 8197–8215.
[15] S.A. Tekalur, A. Shukla, K. Shivakumar, Blast resistance of polyurea based layered composite materials, Compos. Struct. 84 (3) (2008) 271–281.
[16] R.B. Bogoslovov, C.M. Roland, R.M. Gamache, Impact-induced glass transition in elastomeric coatings, Appl. Phys. Lett. 90 (22) (2007) 221910.
[17] N. Iqbal, M. Tripathi, S. Parthasarathy, D. Kumar, P.K. Roy, Polyurea coatings for enhanced blast-mitigation: a review, RSC Adv. 6 (111) (2016) 109706–109717.
[18] M. Grujicic, B.P. d'Entremont, B. Pandurangan, J. Runt, J. Tarter, G. Dillon, Concept-level analysis and design of polyurea for enhanced blast-mitigation performance, J. Mater. Eng. Perform. 21 (10) (2012) 2024–2037.
[19] C.M. Roland, D. Fragiadakis, R.M. Gamache, Elastomer–steel laminate armor, Compos. Struct. 92 (5) (2010) 1059–1064.
[20] D. Fragiadakis, R. Gamache, R.B. Bogoslovov, C.M. Roland, Segmental dynamics of polyurea: effect of stoichiometry, Polymer 51 (1) (2010) 178–184.
[21] K. Sreenivasan, Estimation of hard segment content in polyurethane from solvent absorption, Polym. J. 22 (1990) 620.
[22] N. Iqbal, D. Kumar, P.K. Roy, Understanding the role of isocyanate dilution toward spraying of polyurea, J. Appl. Polym. Sci. 135 (9) (2018). 45869-n/a.
[23] J. White, Origins and measurements of internal stress in plastics, Polym. Test. 4 (2–4) (1984) 165–191.
[24] W. Zoetelief, L. Douven, A. Housz, Residual thermal stresses in injection molded products, Polym. Eng. Sci. 36 (14) (1996) 1886–1896.
[25] I. Yilgor, E. Yilgor, I.G. Guler, T.C. Ward, G.L. Wilkes, FTIR investigation of the influence of diisocyanate symmetry on the morphology development in model segmented polyurethanes, Polymer 47 (11) (2006) 4105–4114.
[26] I. Yilgor, E. Yilgor, S. Das, G. Wilkes, Time-dependent morphology development in segmented polyetherurea copolymers based on aromatic diisocyanates, J. Polym. Sci. B 47 (5) (2009) 471–483.
[27] Y. He, D. Xie, X. Zhang, The structure, microphase-separated morphology, and property of polyurethanes and polyureas, J. Mater. Sci. 49 (21) (2014) 7339–7352.
[28] T. Choi, D. Fragiadakis, C.M. Roland, J. Runt, Microstructure and segmental dynamics of polyurea under uniaxial deformation, Macromolecules 45 (8) (2012) 3581–3589.
[29] M. Grujicic, B. Pandurangan, W.C. Bell, B.A. Cheeseman, C.F. Yen, C.L. Randow, Molecular-level simulations of shock generation and propagation in polyurea, Mater. Sci. Eng. A 528 (10–11) (2011) 3799–3808.
[30] J. LeBlanc, N. Gardner, A. Shukla, Effect of polyurea coatings on the response of curved E-Glass/Vinyl ester composite panels to underwater explosive loading, Compos. Part B. Eng. 44 (2013) 565–574.
[31] Z. Jia, A.V. Amirkhizi, W. Nantasetphong, S. Nemat-Nasser, Experimentally-based relaxation modulus of polyurea and its composites, Mech. Time. Depend. Mater. 20 (2) (2016) 155–174.
[32] W. Huang, Y. Yang, H. Li, P. Lyu, R. Zhang, Characterization and damping property of a modified polyurea material, DEStech. Trans. Mater. Sci. Eng. (ictim) (2017).
[33] J. Qiao, A.V. Amirkhizi, K. Schaaf, S. Nemat-Nasser, G. Wu, Dynamic mechanical and ultrasonic properties of polyurea, Mech. Mater. 43 (10) (2011) 598–607.
[34] F. Ghezzo, X.G. Miao, C.L. Ji, R.P. Liu, On the damping properties of a polyurea elastomer, Adv. Mater. Res., Trans Tech Publ (2014) 21–25.
[35] L. Ning, W. De-Ning, Y. Sheng-Kang, Hydrogen-bonding properties of segmented polyether poly (urethane urea) copolymer, Macromolecules 30 (15) (1997) 4405–4409.
[36] M. Grujicic, J.S. Snipes, S. Ramaswami, R. Yavari, M.K. Ramasubramanian, Meso-scale computational investigation of shock-wave attenuation by trailing release wave in different grades of polyurea, J. Mater. Eng. Perform. 23 (2014) 49.
[37] O. Al-Amoudi, M. Maslehuddin, M. Shameem, M. Ibrahim, Shrinkage of plain and silica fume cement concrete under hot weather, Cem. Concr. Compos. 29 (9) (2007) 690–699.
[38] C. Mazzotti, A. Bilotta, C. Carloni, F. Ceroni, T. D'Antino, E. Nigro, C. Pellegrino, Bond between EBR FRP and concrete, in: C. Pellegrino, J. Sena-Cruz (Eds.), Design Procedures for the Use of Composites in Strengthening of Reinforced Concrete Structures: State-of-the-Art Report of the RILEM Technical Committee 234-DUC, Springer, Netherlands, Dordrecht, 2016, pp. 39–96.
[39] M. Horgnies, P. Willieme, O. Gabet, Influence of the surface properties of concrete on the adhesion of coating: characterization of the interface by peel test and FT-IR spectroscopy, Prog. Org. Coat. 72 (3) (2011) 360–379.

Review of Heterocyclic Scaffolds for the Inhibitors of ATP synthase

Saurabh Mehta*
Department of Applied Chemistry, Delhi Technological University, Bawana Road, Delhi, 110042 India
*Corresponding Author E-mail: **saurabh.mehta@dtu.ac.in; saurabh.dtu@gmail.com**

ABSTRACT:
ATP synthase is an important enzyme and is an established drug target. The involvement of this enzyme has been attributed to various serious diseases and physiological conditions, such as cancer, obesity, neurodegenerative disorders, *etc*. It makes it an attractive drug target whose activity may be modulated by small molecules. A variety of inhibitors reported in the literature have been reviewed, and the important *O*- and *N*- containing heterocyclic scaffolds have been identified and are presented here. On examining the scaffolds, it is observed that a variety of 5-membered ring containing compounds to macrocycles inhibit ATP synthase. These include Flavones and Isoflavons, Catechins, Estrogen metabolites, Polyenic α-pyrones, Amphiphilic cationic dyes, Tertiary amine local anesthetics (TALAs), *N*-sulfonyl tetrahydrobenzodiazepines, *N*-[1-Aryl-2-(1-imidazolo)ethyl]-acylguanidine derivatives, *O*-[1-Aryl-2-(1-imidazolo)ethyl]-thiourethane derivatives and 4-amino benzopyrans.

KEYWORDS: ATP synthase, Drug Discovery, Enzyme, Heterocycle, Inhibitor..

INTRODUCTION:

ATP synthase (F_1F_0), also known as ATP phosphohydrolase (H^+-transporting), is an important enzyme and a complex biomacromolecule. In most of the living organisms including the humans, the majority of ATP is synthesized from ADP and phosphate by the ATP synthase enzyme, using a transmembrane proton motive force, or sodium motive force as the source of energy.[1] The research work on this important enzyme resulted in the award of 1997 Nobel Prize in Chemistry to Professor Paul D. Boyer, Dr. John E. Walker and Professor Jens C. Skou.[2] ATP synthase is a membrane enzyme and is located in the membranes of chloroplast, bacteria, and mitochondria, where it is involved in energy transduction. From a structural viewpoint, the ATP synthase is composed of two rotary motors, the membrane-embedded F0 motor that translocates the coupling ions and the extrinsic F1 motor that synthesizes the ATP.[3] The molecular machinery of ATP-synthase has been previously studied in detail.[4] Furthermore, its catalytic and mechanic cycles have also been explored and reported before.[5,6] There is evidence about ATP synthase being associated with several human ailments. The examples include a neurodegenerative disease namely, Leigh syndrome which causes a neuromuscular disorder.[7] Another neurodegenerative condition, Alzheimer's disease, has been related to a deficiency of ATP synthase in mitochondria.[8] Moreover, the ATP synthase has also been associated with the angiogenesis process in the process of tumor growth.[9] Due to its involvement in various important physiological mechanisms, ATP synthase has emerged as an important drug target.[4,10] It has been successfully demonstrated[11,12] that the inhibition of the proton pump of Mtb ATP synthase using the drug

Bedaquiline (Figure 1) can be used as a therapeutic strategy for the treatment of tuberculosis. Another drug, a mitochondrial ATP synthase inhibitor, 1,4-benzodiazepine, **Bz-423**, was developed for therapy of an autoimmune disorder systemic lupus erythematosus.[9]

Figure 1. Drugs known to inhibit ATP-synthase

Role of rotational mechanism has been studied using inhibitors and mechanisms.[13,14] Furthermore, various types of synthetic as well as natural inhibitors have been studied for ATP Synthase.[15,16,17] Moreover, it has been shown that during myocardial ischemia, the ATP synthase is switched to ATP hydrolase.[18] It has been an established fact that the serious diseases like cancer can be treated using natural and synthetic molecules.[19] Moreover, due to our continuous interest in the synthesis and biological evaluation of important heterocyclic compounds, we became interested in the biological profile of this important drug target.[20,21] Here various classes of compounds/scaffolds for the inhibition of ATP synthase and the related ATP hydrolase enzymes are discussed. Nevertheless, the discussion has been limited to the small molecules and heterocyclic compounds with Mole. Wt. <500, and therefore biomolecules, peptides, nucleotides, *etc*. have been skipped in this study.

DISCUSSION:

Various classes of heterocyclic compounds have been explored against the ATP synthase. The generic structures/scaffolds of the inhibitors have been presented in the Table 1. Flavones and isoflavones are polyphenolic compounds and differ in the position of the phenyl group on the benzopyrone core. Flavone derivative quercetin inhibits the ATPase activities of F0 F1 (Table 1, **A**).[22] On the other hand, Genistein inhibits both the ATP hydrolysis and ATP synthesis activities of mitochondrial ATP synthase, presumably by targeting F0 motor.[23] Among the Catechins [also known as flavan-3-ols (Table 1, **B**)], epicatechin gallate and epigallocatechin gallate inhibit the ATP hydrolysis activity of ATP synthase. Epigallocatechin gallate, exhibits about three times higher potency than epicatechin gallate for the inhibition of ATPase.[23]

The micoestrogens, α-zearalenol and β-zearalenol (Table 1, **C**), are known to inhibit mitochondrial F_0F_1-ATPase enzyme activity. The I_{50} value of α-zearalenol is about 50 μM, and the inhibitory potency of α-zearalenol is 3-4 times greater than that of β-zearalenol. While the mechanism of inhibition by the estrogen metabolites is not clear, the ATP synthase oligomycin sensitivity conferring protein (OSCP) subunit has been identified as an estradiol binding protein, and it has been reported that the inhibition is mediated by the binding of estrogens to this subunit.[24]

α-Pyrone (Table 1, **D**) is a heterocyle containing a six-membered lactone ring. Its derivatives are prevalent in the nature, and some α-pyrone-containing mycotoxins, such as Aurovertin and Citreoviridin, are known to inhibit ATP synthase by targeting its F1 motor. Citreoviridin, is an α-Pyrone derivative, is a natural product, and it inhibits the ATPase activities of F1 from bacteria and mitochondria by binding to the β–subunit of ATP synthase.[25,26]

Amphiphilic cationic dyes like Rhodamine 123 (Table 1, **E**) contain an amine group and another portion that is lipophilic in nature. These dyes are known to inhibit the ATPase activities. Many of these Rhodamines are considered to bind the target at more than one binding sites.[27]

Table 1. Privileged Heterocyclic Scaffolds for the Inhibition of ATP synthase.

TALAs (Tertiary amine local anesthetics), *e.g.* Chlorpromazine and trifluoroperazine (Table 1, **F**), are made up of an aromatic portion, an intermediate chain, and an amine group at the terminal. The TALAs are known to inhibit bacterial ATP synthases selectively.[28,29] Furthermore, the binding sites may differ for different TALA derivatives. In some cases it has been found that the inhibitor forms covalent bonds with the ATP synthase, leading to its irreversible inhibition.[30] *N*-Sulfonyl-substituted tetrahydrobenzodiazepine analogues (Table 1, **G**) are also known to inhibit ATP synthase.[31] 4-amino benzopyran derivatives (Table 1, **H**) are inhibitors of ATP hydrolysis of mitochondrial ATP synthase.[32] It was found that both the amino and benzopyran portions are necessary for inhibition. Other inhibitors of ATP synthase include Cyano- and acylguanidine derivatives (Table 1, **I**) containing imidazolo-ethyl and aryl groups and *O*-[1-Aryl-2-(1-imidazolo)ethyl]-thiourethane (Table 1, **J**) derivatives.[33]

CONCLUSIONS:

Thus a variety of *N*- and *O*- containing heterocyclic scaffolds have been reviewed as inhibitors for ATP synthase enzyme. More advancements are expected in this important area of biochemical research.

CONFLICT OF INTEREST:

There is no conflict of interest.

REFERENCES:

(1) Hong, S; Pedersen, PL. ATP Synthase and the Actions of Inhibitors Utilized To Study Its Roles in Human Health, Disease, and Other Scientific Areas. Microbiology and Molecular Biology Reviews 2008, 72 (4), 590–641.

(2) Press Release: The 1997 Nobel Prize in Chemistry. https://www.nobelprize.org/nobel_prizes/chemistry/laureates/1997/press.html.

(3) Capaldi, RA, Aggeler, R. Mechanism of the F1F0-Type ATP Synthase, a Biological Rotary Motor. Trends in Biochemical Sciences 2002, 27 (3), 154–160.

(4) Boyer, PD. The ATP Synthase—A Splendid Molecular Machine. Annual Review of Biochemistry 1997, 66 (1), 717–749.

(5) Dimroth P, von Ballmoos C, Meier T. Catalytic and Mechanical Cycles in F-ATP Synthases: Fourth in the Cycles Review Series. EMBO reports 2006, 7 (3), 276–282.

(6) Nakamoto RK, Baylis Scanlon JA, Al-Shawi MK. The Rotary Mechanism of the ATP Synthase. Archives of Biochemistry and Biophysics 2008, 476, 43–50.

(7) de Vries DD, et al. A Second Missense Mutation in the Mitochondrial ATPase 6 Gene in Leigh's Syndrome. Annals of Neurology 1993, 34 (3), 410–412.

(8) Schägger H, Ohm TG. Human Diseases with Defects in Oxidative Phosphorylation. European Journal of Biochemistry 2008, 227 (3), 916–921.

(9) Moser TL, et al. Endothelial Cell Surface F1-F0 ATP Synthase Is Active in ATP Synthesis and Is Inhibited by Angiostatin. Proceedings of the National Academy of Sciences of the United States of America 2001, 98 (12), 6656–6661.

(10) Capaldi RA, Aggeler R. Mechanism of the F1F0-Type ATP Synthase, a Biological Rotary Motor. Trends in Biochemical Sciences 2002, 27 (3), 154–160.
(11) Andries K et al. A Diarylquinoline Drug Active on the ATP Synthase of Mycobacterium Tuberculosis. Science (New York, N.Y.) 2005, 307 (5707), 223–227.
(12) Huitric E et al. In Vitro Antimycobacterial Spectrum of a Diarylquinoline ATP Synthase Inhibitor. Antimicrobial agents and chemotherapy 2007, 51 (11), 4202–4204.
(13) Sekiya M, et al. Role of *a*/*b* Interface in F1 ATPase Rotational Catalysis Probed by Inhibitors and Mutations. International Journal of Biological Macromolecules 2017, 99, 615–621.
(14) Nakanishi-Matsui M, Sekiya M, Futai, M. ATP Synthase from Escherichia Coli: Mechanism of Rotational Catalysis, and Inhibition with the ε Subunit and Phytopolyphenols. Biochimica et Biophysica Acta 2016, 1857, 129–140.
(15) Sekiya M, et al. Mayumi Nakanishi-Matsui a. A Unique Mechanism of Curcumin Inhibition on F1 ATPase. Biochemical and Biophysical Research Communications 2014, 452, 940–944.
(16) Lebowitz MS, Pedersen PL. Protein Inhibitor of Mitochondrial ATP Synthase: Relationship of Inhibitor Structure to pH-Dependent Regulation. Archives of Biochemistry and Biophysics 1996, 330 (2), 342–354.
(17) Moreno AJM et al. Mechanism of Inhibition of Mitochondrial ATP Synthase by 17β−Estradiol. Journal of Bioenergetics and Biomembranes 2013, 45 (3), 261–270.
(18) Grover GJ, Malm, J. Pharmacological Profile of the Selective Mitochondrial F1F0 ATP Hydrolase Inhibitor BMS-199264 in Myocardial Ischemia. Cardiovascular Therapeutics 2008, 26 (4), 287–296.
(19) Sharma B. Synthetic and Natural Compounds as Anti-Cancer Agents – A Review. Asian Journal of Research in Chemistry 2017, 10 (5), 699–707.
(20) Allen BK et al. Large-Scale Computational Screening Identifies First in Class Multitarget Inhibitor of EGFR Kinase and BRD4. Scientific Reports 2015, 5, 16924.
(21) Huang Q et al. Design, Synthesis, and Evaluation of NO-Donor Containing Carbonic Anhydrase Inhibitors To Lower Intraocular Pressure. J. Med. Chem. 2015, 58 (6), 2821–2833.
(22) Lang DR, Racker E. Effects of Quercetin and F1 Inhibitor on Mitochondrial ATPase and Energy-Linked Reactions in Submitochondrial Particles. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1974, 333 (2), 180–186.
(23) Zheng J, Ramirez VD. Inhibition of Mitochondrial Proton F0F1-ATPase/ATP Synthase by Polyphenolic Phytochemicals. British Journal of Pharmacology 2000, 130 (5), 1115–1123.
(24) Zheng J, Ramirez VD. Purification and Identification of an Estrogen Binding Protein from Rat Brain: Oligomycin Sensitivity-Conferring Protein (OSCP), a Subunit of Mitochondrial FOF1-ATP synthase/ATPase. Journal of Steroid Biochemistry and Molecular Biology 1999, 68 (1–2), 65–75.
(25) Gause E, Buck M, Douglas M. Binding of Citreoviridin to the Beta Subunit of the Yeast F1-ATPase. J Biol Chem. 1981, 256 (2), 557–559.
(26) Satre M, Bof M, Vignais P. Interaction of Escherichia Coli Adenosine Triphosphatase with Aurovertin and Citreoviridin: Inhibition and Fluorescence Studies. J. Bacteriol. 1980, 142, 768–776.
(27) Bullough DA et al. Inhibition of the Bovine-Heart Mitochondrial F1-ATPase by Cationic Dyes and Amphipathic Peptides. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1989, 975 (3), 377–383.
(28) Chazotte B, Vanderkooi G, Chignell D. Further Studies on F1-ATPase Inhibition by Local Anesthetics. BBA - Bioenergetics 1982, 680 (3), 310–316.
(29) Vanderkooi G et al. On the Mechanism of Action of Anesthetics. Direct Inhibition of Mitochondrial F1-ATPase by N-Butanol and Tetracaine. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1981, 635 (1), 200–203.
(30) Dabbeni-Sala, F, Schiavo G, Palatini P. Mechanism of Local Anesthetic Effect on Mitochondrial ATP Synthase as Deduced from Photolabelling and Inhibition Studies with Phenothiazine Derivatives. Biochimica et Biophysica Acta (BBA) - Biomembranes 1990, 1026, 117–125.
(31) Hamann LG at al. Benzodiazepine-Based Selective Inhibitors of Mitochondrial F1F0 ATP Hydrolase. Bioorganic and Medicinal Chemistry Letters 2004, 14, 1031–1034.
(32) Atwal KS et al. Small Molecule Mitochondrial F 1F 0 ATPase Hydrolase Inhibitors as Cardioprotective Agents. Identification of 4-(N-Arylimidazole)-Substituted Benzopyran Derivatives as Selective Hydrolase Inhibitors. Journal of Medicinal Chemistry 2004, 47 (5), 1081–1084.
(33) Atwal KS et al. N-[1-Aryl-2-(1-Imidazolo)ethyl]-Guanidine Derivatives as Potent Inhibitors of the Bovine Mitochondrial F1F0 ATP Hydrolase. Bioorganic and Medicinal Chemistry Letters 2004, 14, 1027–1030.

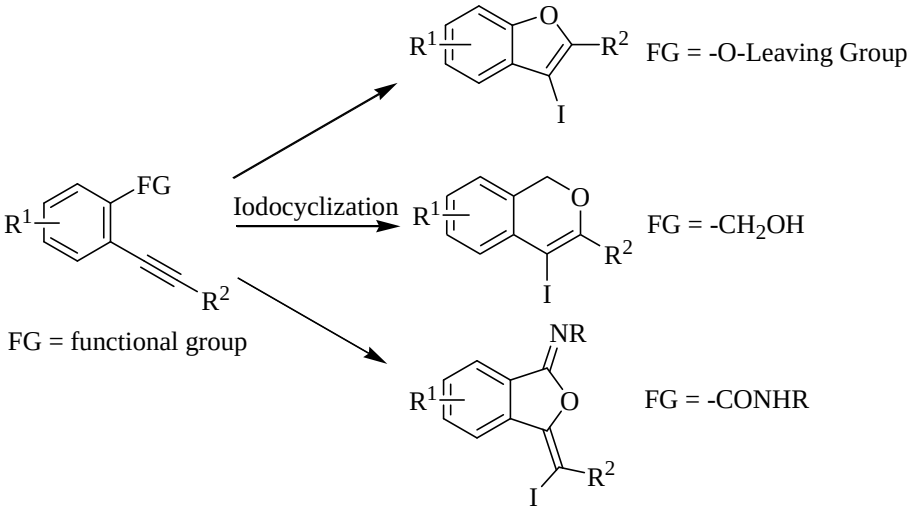
Received on 25.12.2017 Modified on 13.01.2018
Accepted on 06.02.2018 © AJRC All right reserved
Asian J. Research Chem. 2018; 11(2):505-508.
DOI: 10.5958/0974-4150.2018.00090.1

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Review of the Progress for the Synthesis of *N*- and *O*-Heterocycles through Iodine-mediated Cyclizations

Saurabh Mehta*
Department of Applied Chemistry,
Delhi Technological University, Bawana Road, Delhi, India

ABSTRACT



In this review, the recent progress with respect to synthesis of a few important nitrogen and oxygen heterocycles through the iodocyclization strategy has been briefly summarized. The alkyne substrates containing nucleophilic functional groups can be efficiently cyclized to afford important heterocycles with diverse substitution patterns. In general there is good control over the stereo- and regionselectivity of the cyclization products. The cyclization products can be diversified through metal-catalyzed coupling reactions.

INTRODUCTION

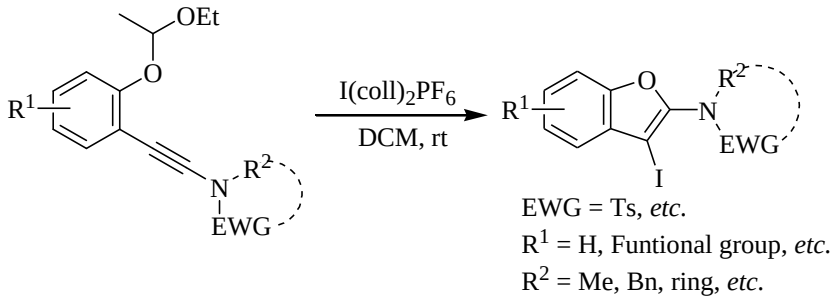
Heterocycles compounds are very important in the areas of pharmaceuticals, agrochemicals, materials research, etc. Several privileged heterocyclic cores such as benzofurans,[1,2] indolinones,[3–5] *etc*. are frequently observed in natural products as well as other biologically important compounds. Synthetic and medicinal chemists put great efforts for developing novel syntheses and compounds for various applications. In the recent past many research group around the world have developed interesting organic syntheses of important heterocyclic and carbocyclic ring systems. One of the popular ways for making these privileged scaffolds is through the electrophilic cyclization, particularly iodocyclization.[6,7] Here we present a review for the recent literature for the synthesis of a few important nitrogen and oxygen containing heterocycles such as Benzo[*b*]furans, Isochromenes,[8,9] Iminolactones,[10] and other fused compounds containing multiple heterocyclic rings.[11,12] It is notable that these scaffolds are very important and their importance is well established in the various areas including medicinal chemistry.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

DISCUSSION

An efficient synthesis of Benzofurans *via* the iodocyclization of the ethoxyethyl ether-protected substrates was reported by Okitsu *et al.* (Scheme 1).[13] An interesting feature of the methodology was that the reactions are performed under mild conditions and get over within seconds affording high yields of the desired products. A similar synthesis was used for the preparation of an antiarrhythmic drug Dronedarone.[14]

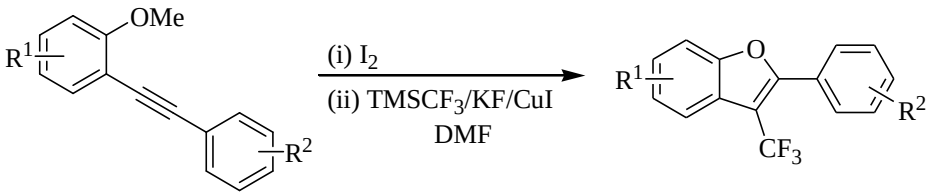
Scheme 1.



EWG = Ts, *etc.*
R^1 = H, Funtional group, *etc.*
R^2 = Me, Bn, ring, *etc.*

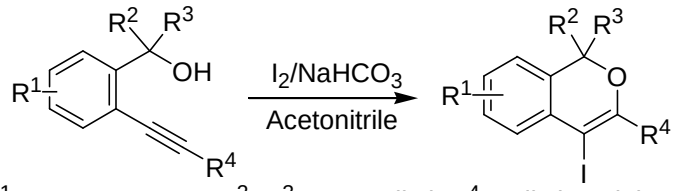
An interesting synthesis of 3-trifluoromethylbenzofurans was reported by Wang *et al.*[15] The one-pot tandem iodocyclization and trifluoromethylation results in the desired products in good yields (Scheme 2).

Scheme 2.



We also reported a mild synthesis of iodine containing isochromene derivatives via the iodocyclization of Alkynyl benzylic alcohol substrates (Scheme 3).[16]

Scheme 3. Preparation of Isochromenes *via* iodocyclization of Alkynyl benzylic alcohols.

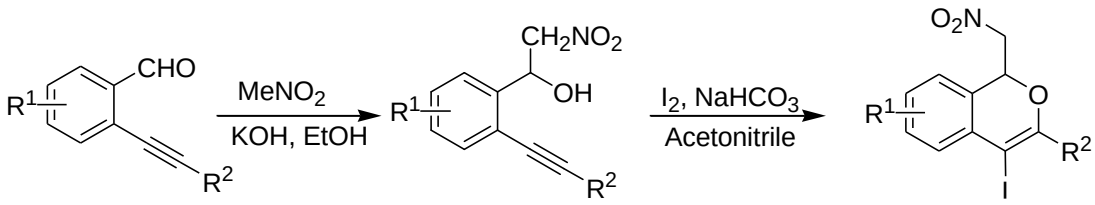


R^1 = H, OMe, NO_2; R^2, R^3 = H or alkyl; R^4 = alkyl, aryl, heteroaryl

A similar iodocyclization strategy was reported by Kundu et al. (Scheme 4).[17] The alkynyl aldehyde substrates were first converted to the corresponding alcohols through the reaction with external nucleophile nitromethane, and the alcohols were eventually iodocyclized affording the Isochromenes.

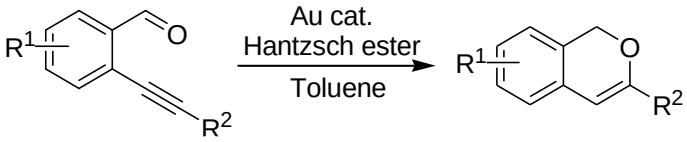
International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Scheme 4. 2-step synthesis of Isochromene derivatives from 2-Alkynyl benzaldehydes.



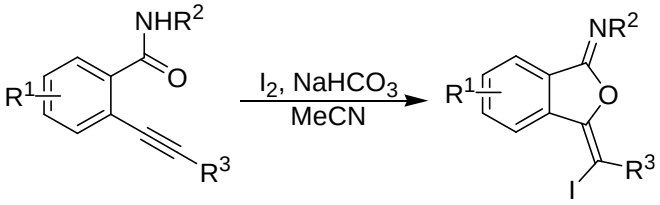
An interesting Synthesis of 1*H*-Isochromenes involving a Gold-catalyzed Domino Cycloisomerization/Reduction strategy was reported by Michelet et al. (Scheme 5).[18]

Scheme 5. Au-catalyzed Cycloisomerization/Reduction Methodology.



We also studied the iodocyclization of Alkynyl benzamide substrates and reported that the reaction under mildly basic conditions leads to the synthesis of iminolactones/Cyclic imidates (Scheme 2).[19] The stereochemistry of the products was confirmed NMR spectroscopy as well as single crystal X-ray crystallographic evidence.

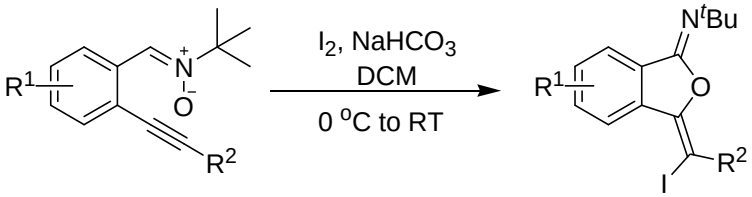
Scheme 6. Regio- and Stereoselective synthesis of Cyclic Imidates *via* Iodocyclization of Alkynylbenzamides



***E*-products**

The cyclization of Nitrone-containing alkyne substrates also resulted in the the formation of corresponding iminolactones (Scheme 7).[20]

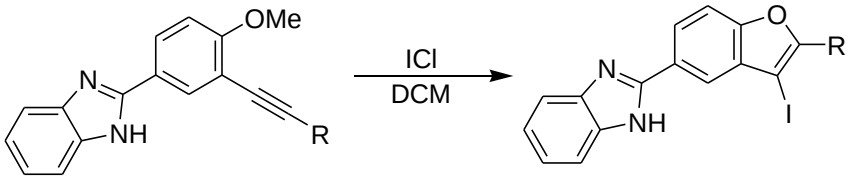
Scheme 7: Iodocyclization of Alkynyl Nitrones



The synthesis of Polyheterocyclic compounds through the Palladium- and Iodine mediated cyclization reactions of appropriately substituted alkynes has been reported by us in the past.[21] It was observed that these reactions may be used to generate linked as well as fused systems very conveniently. Soodamani et al. Recently reported an interesting synthesis of novel compounds containing benzofuranyl and benzimidazole scaffolds linked together (Scheme 8).[22]

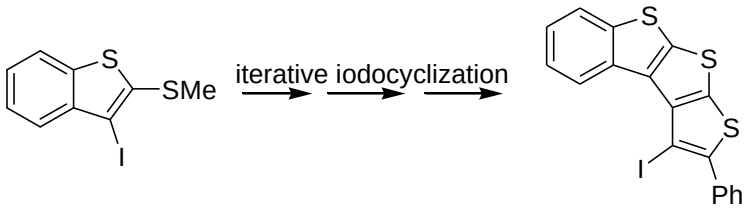
International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 5
May 2018

Scheme 8. Synthesis of benzofuranyl benzimidazoles



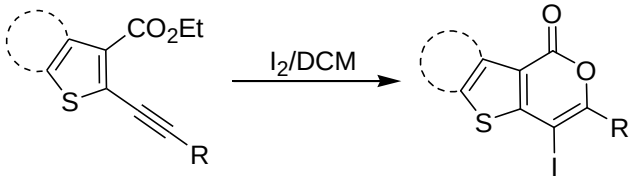
Flynn et al. also reported an interesting synthesis of fused thiophenes through iterative iodocyclizations (Scheme 9).[23]

Scheme 9. Access to Polyfused thiophenes



Pal et al. used an interesting iodocyclization strategy for the synthesis of novel thieno[3,2-c]pyran-4-one derivatives (Scheme 10),[24] and also evaluated their anticancer activities. A few of these derivatives showed the Inhibitory activity in a some cancer cell lines.

Scheme 10. Synthesis of fused hetrocycles such as thieno[3,2-c]pyran-4-ones.



CONCLUSIONS

Thus, various important heterocyclic scaffolds such as benzofurans, iminolactones, ishoochromenes, such as can be efficiently synthesized through the iodocyclization (using I_2/ICl) of alkyne substrates. The reactions can be performed under mild conditions in general and the yields are also good. It has also been demonstrated that these reactions can be futher designed to afford interesting linked and fused heterocyclic ring systems.

REFERENCES

(1) Khanam, H.; Shamsuzzaman. Bioactive Benzofuran Derivatives: A Review. *Eur. J. Med. Chem.* **2015**, *97* (1), 483–504.

(2) Okitsu, T.; Ogasahara, M.; Wada, A. Convergent Synthesis of Dronedarone, an Antiarrhythmic Agent. *Chem. Pharm. Bull.* **2016**, *64* (8), 1149–1153.

(3) Yoo, K.-D.; Park, E.-S.; Lim, Y.; Kang, S.-I.; Yoo, S.-H.; Won, H.-H.; Kim, Y.-H.; Yoo, I.-D.; Yoo, H.-S.; Hong, J. T.; et al. Clitocybin A, a Novel Isoindolinone, From the Mushroom Clitocybe Aurantiaca, Inhibits Cell Proliferation Through G1 Phase Arrest by Regulating the PI3K/Akt Cascade in Vascular Smooth Muscle Cells. *Journal of Pharmacological Sciences* **2012**, *118* (2), 171–177.

(4) Ito, S.; Hirata, Y.; Nagatomi, Y.; Satoh, A.; Suzuki, G.; Kimura, T.; Satow, A.; Maehara, S.; Hikichi, H.; Hata, M.; et al. Discovery and Biological Profile of Isoindolinone Derivatives as Novel Metabotropic Glutamate Receptor 1 Antagonists: A Potential Treatment for Psychotic Disorders. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19* (18), 5310–5313.

(5) Riedinger, C.; Endicott, J. A.; Kemp, S. J.; Smyth, L. A.; Watson, A.; Valeur, E.; Golding, B. T.; Griffin, R. J.; Hardcastle, I. R.; Noble, M. E.; et al. Analysis of Chemical Shift Changes Reveals the Binding Modes of Isoindolinone Inhibitors of the MDM2-p53 Interaction. *Journal of the American Chemical Society* **2008**, *130* (47), 16038–16044.

(6) Godoi, B.; Schumacher, R. F.; Zeni, G. Synthesis of Heterocycles via Electrophilic Cyclization of Alkynes

Containing Heteroatom. *Chemical Reviews*. American Chemical Society 2011, pp 2937–2980.
(7) Dubrovskiy, A. V; Larock, N. A. M. and R. C. Iodocyclization, Followed by Palladium-Catalyzed Coupling: A Versatile Strategy for Heterocyclic Library Construction. *Combinatorial Chemistry & High Throughput Screening* **2012**, *15* (6), 451–472.
(8) Gao, J.-M.; Yang, S.-X.; Qin, J.-C. Azaphilones: Chemistry and Biology. *Chem. Rev.* **2013**, *113*, 4755–4811.
(9) Kanokmedhakul, S.; Kanokmedhakul, K.; Nasomjai, P.; Louangsysouphanh, S.; Soytong, K.; Isobe, M.; Kongsaeree, P.; Prabpai, S.; SuSomdej, A.; Kwanjai, K.; et al. Antifungal Azaphilones from the Fungus Chaetomium Cupreum CC3003. *J. Nat. Prod.* **2006**, *69*, 891–895.
(10) Royer, R. E.; Deck, L. M.; Vander Jagt, T. J.; Martinez, F. J.; Mills, R. G.; Young, S. A.; Vander Jagt, D. L. Synthesis and Anti-HIV Activity of 1,1'-Dideoxygossypol and Related Compounds. *Journal of Medicinal Chemistry* **1995**, *38* (13), 2427–2432.
(11) Audoux, J.; Achelle, S.; Turck, A.; Marsais, F.; Plé, N. Synthesis of New Flat Polyheterocyclic Systems Potential DNA Intercalating Agents Diazines Part 47. *Journal of Heterocyclic Chemistry* **2006**, *43* (6), 1497–1503.
(12) Mcculloch, I.; Heeney, M.; Chabinyc, M. L.; Delongchamp, D.; Kline, R. J.; Cölle, M.; Duffy, W.; Fischer, D.; Gundlach, D.; Hamadani, B.; et al. Semiconducting Thienothiophene Copolymers: Design, Synthesis, Morphology, and Performance in Thin-Film Organic Transistors. *Advanced Materials* **2009**, *21* (10–11), 1091–1109.
(13) Okitsu, T.; Nakata, K.; Nishigaki, K.; Michioka, N.; Karatani, M.; Wada, A. Iodocyclization of Ethoxyethyl Ethers to Ynamides: An Immediate Construction to Benzo[b]furans. *J. Org. Chem.* **2014**, *79* (12), 5914–5920.
(14) Okitsu, T.; Nakazawa, D.; Taniguchi, R.; Wada, A. Iodocyclization of Ethoxyethyl Ethers to Alkynes: A Broadly Applicable Synthesis of 3-Iodobenzo[b]furans. *Org. Lett.* **2008**, *10* (21), 4967–4970.
(15) Wang, W. Y.; Hu, B. L.; Deng, C. L.; Zhang, X. G. One-Pot Synthesis of 3-Trifluoromethylbenzofurans via Tandem Iodocyclization and Trifluoromethylation of 2-Alkynylanisoles. *Tetrahedron Letters* **2014**, *55* (8), 1501–1503.
(16) Mancuso, R.; Mehta, S.; Gabriele, B.; Salerno, G.; Jenks, W. S.; Larock, R. C. A Simple and Mild Synthesis of 1H-Isochromenes and (Z)-1-Alkylidene-1,3-Dihydroisobenzofurans by the Iodocyclization of 2-(1-Alkynyl)benzylic Alcohols. *J. Org. Chem.* **2010**, *75* (3), 897–901.
(17) Arigela, R. K.; Samala, S.; Mahar, R.; Shukla, S. K.; Kundu, B. Synthesis of Triazolo Isoquinolines and Isochromenes from 2-Alkynylbenzaldehyde via Domino Reactions under Transition-Metal-Free Conditions. *J. Org. Chem.* **2013**, *78*, 10476–10484.
(18) Tomás-Mendivil, E.; Starck, J.; Ortuno, J.-C.; Michelet, V. Synthesis of Functionalized 1H-Isochromene Derivatives via a Au-Catalyzed Domino Cycloisomerization/Reduction Approach. *Org. Lett.* **2015**, *17*, 6126–6129.
(19) Mehta, S.; Yao, T.; Larock, R. C. Regio- and Stereoselective Synthesis of Cyclic Imidates via Electrophilic Cyclization of 2-(1-Alkynyl)benzamides. A Correction. *J. Org. Chem.* **2012**, *77* (23), 10938–10944.
(20) Chen, D.; Song, G.; Jia, A.; Li, X. Gold- and Iodine-Mediated Internal Oxygen Transfer of Nitrone- and Sulfoxide-Functionalized Alkynes. *The Journal of Organic Chemistry* **2011**, *76* (20), 8488–8494.
(21) Mehta, S.; Larock, R. C. Iodine/palladium Approaches to the Synthesis of Polyheterocyclic Compounds. *J. Org. Chem.* **2010**, *75* (5), 1652–1658.
(22) Soodamani, V.; Patel, D.; Nayakanti, D.; Josyula, R. A Facile Synthesis of Novel Benzofuranyl Benzimidazole Derivatives. *Journal of Heterocyclic Chemistry* **2014**, *51* (5), 1491–1499.
(23) Aurelio, L.; Volpe, R.; Halim, R.; Scammells, P. J.; Flynn, B. L. Synthesis of Thieno-Fused Heterocycles through Reiterative Iodocyclization. *Advanced Synthesis & Catalysis* **2014**, *356* (9), 1974–1978.
(24) Nakhi, A.; Adepu, R.; Rambabu, D.; Kishore, R.; Vanaja, G. R.; Kalle, A. M.; Pal, M. Thieno[3,2-C]pyran-4-One Based Novel Small Molecules: Their Synthesis, Crystal Structure Analysis and in Vitro Evaluation as Potential Anticancer Agents. *Bioorganic & Medicinal Chemistry Letters* **2012**, *22* (13), 4418–4427.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/324586067

Review on Effects of different parameters on Behavior of Geopolymer Concrete

Article · June 2017

CITATION
1

READS
3

1 author:



Manvendra Verma
Delhi Technological University
3 PUBLICATIONS **1** CITATION

SEE PROFILE

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

behavior of Geopolymer concrete View project

All content following this page was uploaded by Manvendra Verma on 18 April 2018.

The user has requested enhancement of the downloaded file.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

Review on Effects of Different Parameters on Behavior of Geopolymer Concrete

Manvendra Verma[1], Nirendra Dev[2]

Ph.D. Student, Department of Civil Engineering, Delhi Technological University, Delhi India[1]

Professor, Department of Civil Engineering, Delhi Technological University, Delhi India[2]

ABSTRACT: Portland cement concrete industry has grown universally in recent years. The demand for concrete as a construction material has increased due to an increase of infrastructure. However, Portland cement concrete generates problems such as durability and carbon dioxide emission. There are many ways to reduce environmental pollution (carbon dioxide) caused by the production of Portland cement and by the increasing of waste material. Around 120 million tons of fly ash get assembled every year at the thermal power stations in India. It becomes a serious problem due to the inadequacy of land disposal. Cement is totally replaced by the pozzolanic material that is rich in Silicon, and Aluminium like fly ash referred to as "Geopolymer concrete" which is a contemporary material. Geopolymer concrete was actually manufactured by reusing and recycling of industrial solid wastes and by-products. Fly Ash, a byproduct of coal obtained from the thermal power plant is plenty available worldwide. Fly ash is used as ingredients in concrete which enhance the properties of concrete and utilization of fly ash is helpful for consumption. This paper presents a brief history and review of geopolymer technology with the aim of introducing the technology and the vast categories of materials that may be synthesized by alkali activation of alumino silicates.

KEYWORDS: Fly Ash, Carbon Dioxide, Alkaline Activator, Geopolymer Concrete, Strength, And Durability.

I. INTRODUCTION

In the context of increased awareness regarding the ill effects of the over-exploitation of natural resources, eco-friendly technologies are to be developed for effective management of these resources. Concrete usage around the world is second only to water. Cement is conventionally used as the primary binder to produce concrete. The environmental problems associated with the production of cement are well known. The amount of the carbon dioxide released during the manufacture of cement due to the calcination of limestone and combustion of fossil fuel is in the order of one ton for every ton of cement produced. Hence, it is imminent to find an alternative material to the existing most expensive most resource consuming Portland cement. OPC is extensively used in India due to its low cost and easy availability. Concrete can be cast in almost any desired shape, and once hardened, can become a structural (load bearing) element. On the other hand, it also affects environment. Also, there are much negative influences of OPC. For example emissions of airborne pollution in the form of gases, noise, dust, and vibration when operating machinery and during blasting in quarries, devouring of large quantities of fuel during manufacture and release of CO2 from the raw materials during manufacture. Due to all such reasons, it is needed to be replaced by non-producing CO2 materials such as fly ash and various supplementary materials. As such, geopolymer concrete has been introduced to reduce this problem. Fly ash is plentifully available to replace totally manufactured cement and make concrete like material. It is an excellent alternative construction material to the existing plain cement concrete. Geopolymer concrete shall be manufactured without using any amount of ordinary Portland cement. The advancement of geopolymer concrete is an important step towards the production of environmentally friendly concrete. Geopolymer is a type of amorphous alumino-silicate cementitious material that exhibits the ideal properties of rock-forming minerals, i.e., hardness, chemical stability, and longevity. Geopolymerization involves a heterogeneous chemical reaction between solid aluminosilicate oxides and alkali metal silicate solutions at highly alkaline conditions and mild

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

temperatures yielding three-dimensional amorphous to semi-crystalline polymeric and ring structures, which consist of Si–O–Al, and Si–O–Si bonds.

II. OBJECTIVE

The main target of this study to analyze the different parameters to affect the geopolymer concrete properties like as durability and strength.

III. LITERATURE REVIEW

Gaurav Nagalia, Yeonho Park, Ali Abolmaali, and Pranesh Aswath examines the role of alkali hydroxide and its concentration on the microstructure and compression strength of fly ash based geopolymer concrete. Geopolymer is an innovative ceramic material, composed of long chains and networks of inorganic molecules, that is used as an alternative to conventional portland cement for civil infrastructure applications. In this study, aluminosilicate geopolymers with different alkaline solutions [NaOH, KOH, $Ba(OH)_2$, and LiOH] were prepared by mixing Class C (9.42% CaO) and Class F-fly ash (1.29% CaO). The samples were cured under different experimental conditions and then tested for compressive strength. X-ray diffraction (XRD) and scanning electron microscopy/energy dispersive spectroscopy (SEM/EDS) have been used to identify the new phases formed in the geopolymeric matrix. The results revealed that microstructural properties of fly ash particles and calcium content have a significant effect on the setting time and the compressive strength of geopolymer concrete, and the structure of the formed Geopolymer is typically glass like. The early polymerization of geopolymer concrete is dependent on the concentration of the sodium hydroxide solution (NaOH). The exclusive use of NaOH showed the highest compressive strength when compared with other combinations of alkali hydroxides.

D. V. Reddy, Jean-Baptiste Edouard and Khaled Sobhan use the supplementary cementitious materials as partial replacements of the cement in concrete will play a significant role in the environmental control of greenhouse effects and global temperature reduction. The development of geopolymer concrete (GPC), in which all of the portland cement is replaced with fly ash (in combination with sodium hydroxide and sodium silicate solutions), offers a promising alternative to ordinary Portland cement concrete. This paper evaluates the durability characteristics of low-calcium fly ash–based geopolymer concrete that was subjected to a corrosive marine environment. A series of GPC beams, which contained fly ash with 8- and 14-M concentrations of NaOH and SiO2=Na2O solutions, respectively, and were centrally reinforced with a 13-mm rebar, were tested for accelerated corrosion exposure with wet and dry cycling in artificial seawater and an induced current. A sudden rise in the current intensity that is attributable to specimen cracking was indicative of durability. The test results indicated the excellent resistance of the geopolymer concrete to chloride attack, with a longer time to corrosion cracking relative to ordinary portland cement concrete.

Cheah Chee Ban, Part Wei Ken and Mahyuddin Raml introduces a novel Geopolymer system based on hybridization between high calcium wood ash (HCWA) and pulverized fuel ash (PFA), which use a low percentage of alkaline activator (5% of binder weight), ambient temperature curing, and a pressurized forming method for the fabrication of load-bearing masonry unit. The elimination of alkali hydroxide was made possible by the Arcanite mineral (approximately 12%), which is inherently present in HCWA. The HCWA-PFA geopolymer mortar blocks with various sodium silicate content (0–5%) and curing regime (water and moist curing) were assessed in terms of mechanical (compressive, flexural, and UPV) and durability performance (water absorption, total porosity, and intrinsic air permeability) over a curing period of up to 90 days. The addition of Na2SiO3 was found to enhance the early strength development of HCWA-PFA mortar blocks. Moreover, the mortar blocks exhibited continuous strength growth over a prolonged curing period. Moist-cured samples consistently outperformed their water-cured counterpart in every aspect. All of the fabricated HCWAPFA geopolymer mortar blocks are classified as normal-weight, solid, load-bearing masonry unit.

Partha Sarathi Deb and Prabir Kumar Sarker presents the effects of ultrafine fly ash on setting time, Strength, and porosity of geopolymer paste specimens cured at 20°C. The addition of ultrafine fly ash reduced

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

setting time and improved compressive strength of geopolymers based on fly ash only or fly ash blended with 15% ground granulated blast-furnace slag (GGBFS) or 10% ordinary portland cement (OPC). With the addition of ultrafine fly ash, a reduction in porosity was observed by the mercury intrusion porosimetry (MIP), and a denser microstructure was observed by the scanning electron microscope (SEM) photographs. Thus, for room temperature curing conditions, a control on the setting time and improvements in strength and porosity of geopolymers could be achieved by varying the percentage of ultrafine fly ash.

Kunal Kupwade-Patil, and Erez N. Allouche The durability of steel-reinforced-concrete specimens made from three alkali-activated fly ash (FA) stockpiles and ordinary portland cement (OPC) in cyclic wet-dry chloride environment was evaluated over a period of 12 months. Testing methods included electrochemical methods, chloride diffusion and contents analysis, chemical and mechanical analyses, and visual examination. Geopolymer concrete (GPC) specimens made from Class F FA exhibited lower diffusion coefficients, chloride contents, and porosity compared with their GPC Class C FA and OPC counterparts. Overall, GPC specimens displayed limited signs of leaching and corrosion product formation, whereas OPC specimens exhibited the formation of multiple corrosion products along with significant leaching.

Zhen Liu, C. S. Cai, Hui Peng and Fenghong Fan Aiming to solve the problems caused by the construction and demolition waste and the depletion of natural aggregates, in the present study, both the coarse recycled aggregates and the fine recycled aggregates are used to produce new green concrete with a fly ash based geopolymer. The mechanical study shows that the reduction rates of the compressive strength, Young's modulus, and the Poisson's ratio of the geopolymeric recycled aggregate concrete (GRAC) specimens increase with the increase of the water/cement (w/c) ratio. The Young's modulus is much more vulnerable to the increase of the w/c ratio than the compressive strength. These significant changes to the GRAC samples could be attributable to the deteriorated properties rendered to the geopolymer by the increase of the w/c ratio. When the minimum ratio of w/c is applied, the fly ash–based geopolymer renders better mechanical properties to the concrete specimens than the ordinary portland cement. With the aid of scanning electron microscopy and nanoindentation, there is no interfacial transition zone well developed between the old cement paste and the new geopolymer/cement paste.

William Gustavo Valencia Saavedra, Daniela Eugenia Angulo, and Ruby Mejía de Gutiérrez evaluated the performance of geopolymer concretes based on a binary mixture of fly ash (FA) with blast furnace slag (GBFS) in an 80=20 ratio and activated with a mixture of sodium silicate and sodium hydroxide. FA/GBFS and Portland cement (OPC) concretes were immersed in 5% by weight sodium sulfate and magnesium sulfate solutions. Volumetric expansion and mechanical resistance loss were measured, and the reaction products were characterized by X-ray diffraction and scanning electron microscopy. The highest levels of deterioration were found in the probes exposed to MgSO4, indicating the higher aggressiveness of this solution. After aging for 360 days in MgSO4, the FA/ GBFS concrete expanded by 0.04% and decreased its mechanical resistance by 33%, compared to 0.08 and 48%, respectively, for similarly treated OPC concretes. However, the same exposure to Na2SO4 increased the observed FA/GBFS resistance while reducing the mechanical resistance of OPC by 30%. This relates to their respective expansion percentages of 0.0068 and 0.412% and the presence of expansive ettringite in the OPC concretes.

Chaicharn Chotetanorm, Prinya Chindaprasirt, Vanchai Sata, Sumrerng Rukzon and Apha Sathonsaowaphak investigated the resistance to sulfate attack, compressive strength, sorptivity, and pore size of high-calcium bottom ash geopolymer mortars. Ground lignite bottom ashes (BAs) with median particle sizes of 16, 25, and 32 μm were used. NaOH, sodium silicate and temperature curing were used to activate the geopolymerization. Results showed that relatively high strengths of 40.0–54.5 Mpa were obtained for the high-calcium bottom ash geopolymer mortars. The use of fine BA improved the strength and resistance to sulfate attack of mortars. The good performances were attributable to the high degree of reaction of fine BA and the associated low amount of large pores (0.05–100 μm) compared with those of coarse BA. The incorporation of water improved the workability of mixes, and the compressive strength, sorptivity, and resistance to sulfate attack decreased due to the increase in large pores.

P. Chindaprasirt, T. Chareerat, S. Hatanaka and T. Cao Investigated the synthesis of high-strength Geopolymer using fine high-calcium fly ash. The effect of fineness of fly ash on the setting time of geopolymer

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

paste, workability, strength development, and drying shrinkage of geopolymer mortars made from classified fine high-calcium fly ash was investigated. Air-classified fly ash with three different finenesses coarse original fly ash (CFA), medium-fineness fly ash (MFA), and fine fly ash (FFA)—from the Mae Moh Power Station was used for this study. The heat-cured geopolymers were activated with sodium hydroxide (NaOH) and sodium silicate. A small amount of water was incorporated for workability. The results indicate that the setting time of paste decreases with an increase in fly-ash fineness. The flow, strength, and drying-shrinkage characteristics of mortars were improved using fine fly ash. Geopolymer mortars with high 28-day compressive strength of 86.0 MPa were obtained.

Kunal Kupwade-Patil, and Erez N. Allouche This study reports the findings of an experimental investigation for alkali silica reaction (ASR) between reactive aggregates and the geopolymer matrix. Specimens were prepared using one Class C and two Class F fly ash stockpiles. Mechanical testing included potential reactivity of the aggregates via length change and compression test measurements, as per ASTM standards. Results suggest that the extent of ASR reaction due to the presence of reactive aggregates in fly ash-based geopolymer concretes is substantially lower than in the case of ordinary portland cement-based concrete, and well below the ASTM specified threshold. Furthermore, geopolymer concrete specimens appeared to undergo a densification process in the presence of alkali solutions, resulting in reduced permeability and increased mechanical strength. Utilizing ASR-vulnerable aggregates in the production of geopolymer concrete products could contribute to the economic appeal and sustainability of geopolymer binders in regions that suffer from the insufficient local supply of high-quality aggregates.

O. Sanusi, B. Tempest, V. O. Ogunro and J. Gergely investigation the effect of replacing some percentage of natural coarse aggregate with recycled concrete aggregate (RCA) in geopolymer concrete synthesis. The study focuses on the compressive strength and release of elements from the Geopolymer concrete samples. The toxicity characteristic leaching procedure (TCLP) was used to determine the release of potentially hazardous heavy metals (PHHMs) and other constituents of potential concern (COPCs) from the geopolymer concrete and also to identify if the material should be categorized as either hazardous or nonhazardous. It was found that replacing virgin aggregates with RCA had a small impact on the compressive strength of the concrete, reducing it by a maximum of 34%. As measured by toxicity characteristics leaching procedure (TCLP), the release of As was reduced by adding RCA. The leached concentrations of the elements, Ba, Cd, Cr, and Pb, were found to be below TCLP regulatory limits.

Pattanapong Topark-Ngarm, Prinya Chindaprasirt, and Vanchai Sata investigated setting time, strength and bond of high-calcium fly ash geopolymer concrete. The high-calcium fly ash was from Mae Moh power plant in northern Thailand. Both sodium silicate solution and sodium hydroxide solution were used as alkali activators in every mix. Sodium hydroxide solution with 10 M, 15 M, and 20 M concentrations, sodium silicate to sodium hydroxide ratios of 1.0 and 2.0, alkaline liquid to fly ash ratio of 0.5 and two curing regimes viz., heat curing at 60 _ 2°C for 24 h and room temperature curing at 23 _ 2°C were used. The results indicated that fresh geopolymer concrete had the short setting time of 28–58 min due to the presence of high calcium content of fly ash. In general, strengths and modulus of elasticity increased with the increase in NaOH concentration. For compressive strength, the optimum Na2O content was around 12% of fly ash. The high-strength geopolymer concrete with a 28-day compressive strength of 54.4 Mpa was obtained for mix with 15M NaOH. The moduli of elasticity of geopolymer concrete were related to the compressive strengths and comparable to those of Portland cement concrete. The tensile splitting strength and bond strength were also related to the compressive strength and the values were higher than those of portland cement concrete. In particular, the bond strengths were significantly higher than those given by the current design code.

Weibo Ren, Jinyu Xu, and Erlei Bai studied the residual behaviors of the alkali-activated fly ash–slag geopolymer concrete (FSGC) under different heating temperatures and cooling regimes. Changes in weight, compressive strength and ultrasonic pulse velocity were firstly investigated. Then the wavelet packet technique was adopted to further analyze the measured ultrasonic signals in frequency domain. Finally, the microstructures were investigated by scanning electron microscopy. The results indicate that an increase in temperature makes the weight, compressive strength, and ultrasonic pulse velocity value of FSGC decrease, especially under extreme heat conditions. The deterioration of specimens cooled by watering is more obvious than that of

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

specimens cooled naturally. Based on the experimental results, it can be concluded that the critical temperature for FSGC is 600°C, but this temperature may change with the contents of fly ash and slag. Also, with the increase of temperature and specimen damage, the higher frequency components of the signal attenuate steadily and signal energy stored in lower frequency bands increases. A wavelet packet energy spectrum-based index is sensitive to thermal damage, which can be used as a supplementary indicator to assess the damage degree of the heated FSGC. After high-temperature exposure, FSGC undergoes distinct morphological destruction as confirmed by the developments of microcracks and the damaged aluminosilicate gel matrix.

Patimapon Sukmak, Pre De Silva, Suksun Horpibulsuk and Prinya Chindaprasirt examines the short-term resistance ability of two different cementitious systems prepared using silty clay as a major component of five wt% sodium sulfate and five wt% magnesium sulfate solutions. The two cementitious systems are clay–Portland cement and clay-high calcium fly ash (FA) geopolymer. The 28-day strength of the clay-FA geopolymer is up to 1.2 times higher than that of claycement. The results show that the physical performance of clay-FA geopolymer, when exposed to sulfate solution, is better than that of claycement. While the geopolymer phase is prominent in the clay-FA geopolymer system, gypsum, and ettringite phases are present in both systems, especially in the clay-cement system. The strength reduction in clay-cement with duration of sulfate exposure is caused by the formation of ettringite and crystallization of gypsum (observed by scanning electron microscope and X-ray diffraction) as well as the decalcification of CSH phases. There is no major change in the microstructure and pH of clay-FA geopolymer when exposed to sulfate solutions, and hence the strength reduction is insignificant. The exposure to magnesium sulfate solution causes more degradation in both clay-FA geopolymer and clay-cement system than the exposure to sodium sulfate because the CSH phases react with magnesium sulfate solution to form ettringite. The disappearance of CSH phases in both systems is observed by XRD analysis. Overall, clay-FA geopolymer shows better resistance to sulfate attack than clay-cement. The clay-FA geopolymer can be considered as a green construction and building material in sulfate-rich areas.

Muhammad Aamer Rafique Bhutta, Warid M. Hussin, Mohd Azreen and Mahmood Mohd Tahir investigated the performance of geopolymer concrete prepared using blended ash of pulverized fuel ash and palm oil fuel ash from agroindustrial waste along with alkaline activators when exposed to a 5% sodium sulphate solution for up to 18 months. Ordinary portland cement (OPC) concrete was also prepared as the control concrete. The main parameters studied were the evaluation of mass change, compressive strength, products of degradation, and microstructural changes. The deterioration was examined using X-ray diffraction (XRD), a Fourier transform infrared spectroscopy (FTIR), a thermogravimetry analyzer (TGA), and a field emission scanning electron microscope (FESEM). The results of geopolymer and OPC concretes were compared and discussed. The compressive strength of geopolymer concrete, when exposed to a 5% sodium sulfate solution for more than one year, was superior to that of OPC concrete. These materials could be used for making sulphate-resistant concretes, attributable to a more stable cross-linked aluminosilicate polymer structure formed in the geopolymer concrete.

Farhad Aslani examine the relationships are proven for normal and high-strength GCs at elevated temperatures to establish efficient modeling and specify the fire-performance criteria for concrete structures. They are developed for unconfined GC specimens that consist of compressive strength, modulus of elasticity, flexural strength, thermal strain, prestressed thermal strain, and the compressive stress–strain relationships at elevated temperatures. The proposed relationships at elevated temperature are compared with experimental results and available OPC relationships. These results are used to establish more accurate and general compressive stress–strain relationships.

IV. CONCLUSION

From the literature review, the critical reviews are as follows:

- Geopolymer concrete use as an alternative to Portland cement and it reduce CO2 emission in the world so it can be said that it helps to the nature that's why geopolymer concrete represents a "Green concrete" and also as an "eco-friendly concrete."

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 6, June 2017

- The experimental works prove that geopolymer concrete with moderate to moderately high strength can be developed at very low concentration of alkaline activator and ambient air curing.
- Use of fly ash in combination with slag is found to be more reactive and the strength gained is more.
- Specialized curing regimes such as stream curing, hot air curing are required only when accelerated strength gain is required. When design strength is required at 28 days normal curing is sufficient.
- Higher the fineness of fly ash gives a higher compressive strength because of more surface area with a more Si-Al bond of polymerization.
- Generally heat cured geopolymer concrete gives higher strength but it can be obtained at ambient tempreture by replacing fly ash content by GGBFS.
- Fly ash based geopolymer is better than normal concrete in many aspects uch as compressive strength, exposure to aggressive environment workability and exposure to high tempreture.

REFERENCES

1. Chaicharn Chotetanorm, Prinya Chindaprasirt, Vanchai Sata, Sumrerng Rukzon and Apha Sathonsaowaphak, "High-Calcium Bottom Ash Geopolymer: Sorptivity, Pore Size, and Resistance to Sodium Sulfate Attack" J. Mater. Civ. Eng., 2013, 25(1): 105-111.
2. P. Chindaprasirt, T. Chareerat, S. Hatanaka and T. Cao, "High-Strength Geopolymer Using Fine High-Calcium Fly Ash" J. Mater. Civ. Eng., 2011, 23(3): 264-270.
3. Kunal Kupwade-Patil and Erez N. Allouche, "Impact of Alkali Silica Reaction on Fly Ash-Based Geopolymer Concrete" J. Mater. Civ. Eng., 2013, 25(1): 131-139.
4. O. Sanusi, B. Tempest, V. O. Ogunro and J. Gergely, "Leaching Characteristics of Geopolymer Cement Concrete Containing Recycled Concrete Aggregates" J. Hazard. Toxic Radioact. Waste, 2016, 20(3): -1—1.
5. Pattanapong Topark-Ngarm, Prinya Chindaprasirt and Vanchai Sata, "Setting Time, Strength, and Bond of High-Calcium Fly Ash Geopolymer Concrete" J. Mater. Civ. Eng., 2015, 27(7): -1—1.
6. Weibo Ren, Jinyu Xu and Erlei Bai, " Strength and Ultrasonic Characteristics of Alkali-Activated Fly Ash-Slag Geopolymer Concrete after Exposure to Elevated Temperatures" J. Mater. Civ. Eng., 2016, 28(2): -1—1.
7. Patimapon Sukmak, Pre De Silva, Suksun Horpibulsuk and Prinya Chindaprasirt, "Sulfate Resistance of Clay-Portland Cement and Clay High-Calcium Fly Ash Geopolymer" J. Mater. Civ. Eng., 2015, 27(5): -1—1.
8. Muhammad Aamer Rafique Bhutta, Warid M. Hussin, Mohd Azreen and Mahmood Mohd Tahir, "Sulphate Resistance of Geopolymer Concrete Prepared from Blended Waste Fuel Ash" J. Mater. Civ. Eng., 2014, 26(11): -1—1.
9. Farhad Aslani, "Thermal Performance Modeling of Geopolymer Concrete" J. Mater. Civ. Eng., 2016, 28(1): -1—1.
10. William Gustavo Valencia Saavedra, Daniela Eugenia Angulo and Ruby Mejía de Gutiérrez, "Fly Ash Slag Geopolymer Concrete: Resistance to Sodium and Magnesium Sulfate Attack" J. Mater. Civ. Eng., 2016, 28(12): -1—1.
11. Zhen Liu, C. S. Cai, Hui Peng and Fenghong Fan, "Experimental Study of the Geopolymeric Recycled Aggregate Concrete" J. Mater. Civ. Eng., 2016, 28(9): -1—1.
12. Kunal Kupwade-Patil and Erez N. Allouche, "Examination of Chloride-Induced Corrosion in Reinforced Geopolymer Concretes" J. Mater. Civ. Eng., 2013, 25(10): 1465-1476.
13. Partha Sarathi Deb and Prabir Kumar Sarker, "Effects of Ultrafine Fly Ash on Setting, Strength, and Porosity of Geopolymers Cured at Room Temperature" J. Mater. Civ. Eng., -1—1.
14. Cheah Chee Ban, Part Wei Ken, and Mahyuddin Ramli, "Effect of Sodium Silicate and Curing Regime on Properties of Load Bearing Geopolymer Mortar Block" J. Mater. Civ. Eng., -1—1.
15. D. V. Reddy, Jean-Baptiste Edouard, and Khaled Sobhan, "Durability of Fly Ash–Based Geopolymer Structural Concrete in the Marine Environment" J. Mater. Civ. Eng., 2013, 25(6): 781-787.
16. Gaurav Nagalia, Yeonho Park, Ali Abolmaali and Pranesh Aswath, "Compressive Strength and Microstructural Properties of Fly Ash–Based Geopolymer Concrete" J. Mater. Civ. Eng., 2016, 28(12): -1—1.

International Journal of Applied Research on Information Technology and Computing
DOI: 10.5958/0975-8089.2018.00007.6

IndianJournals.com
A product of Diva Enterprises Pvt. Ltd.

Segmentation of Dark Foreground Objects by Maximum Non-Extensive Entropy Partitioning

Seba Susan[1]*, Rahul Singh[2], Amit Kumar[3], Abhishek Kumar[4] and Ashwani Kumar[5]

[1]Associate Professor, [2,3,4,5]Undergraduate B.Tech., Department of Information Technology, Delhi Technological University, New Delhi-110042, India

**Corresponding author email id: seba_406@yahoo.in*

ABSTRACT

In this paper, we investigate the segmentation of dark foreground objects in a relatively bright background clutter by maximum non-extensive entropy partitioning of the image histogram. The non-linear, non-extensive entropy with Gaussian gain, which is a regularity indicator, is embedded in the maximum entropy thresholding framework, in the background of the recent work in Susan et al. (2016. Sadhana 41: 1393) on segregating facial intensities related to emotions by iterative maximum entropy partitioning. The non-linearity of this entropy ensures that out of the two partitions of the image histogram, the darkest shades in the image approximate very finely a uniform probability distribution and are separated out effectively from the brighter shades in the image that coarsely approximate a uniform distribution. This definition of the point of maximum entropy brought about by the non-linearity of the Gaussian curve targets and segments out the darkest shades pertaining to the foreground object in a more efficient manner than other thresholding techniques. Comparisons with the existing entropic thresholding schemes on test instances from a benchmark object segmentation dataset confirm the utility and efficiency of our method.

Keywords: Image thresholding, Maximum entropy partitioning, Non-extensive entropy with Gaussian gain, Weighted sum of non-extensive entropies, Entropy-based image thresholding, Maximum non-extensive entropy partitioning, Foreground object detection

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.107 on dated 4-Jun-2018

1. INTRODUCTION

The segmentation of anomalous intensities from images assumes significance in the areas of video surveillance, medical diagnosis and remote sensing[1]. Anomalous objects mostly appear as dark objects against a bright background[2]. Segmenting the foreground object by image thresholding is a risky proposition[3], as the decision regarding each pixel is a binary one relying entirely on the pixel intensity and its placement with respect to the threshold value. On the other hand, image thresholding is the simplest and computationally inexpensive technique for foreground object segmentation. Assuming that in the majority of cases, the foreground object appears uniformly dark against a bright clutter of background objects, we in our work, make use of the non-linear non-extensive entropy with Gaussian gain to compute the threshold that divides the image histogram into two partitions by the maximum entropy principle.

2. IMAGE THRESHOLDING BY MAXIMUM NON-EXTENSIVE ENTROPY PARTITIONING (MNEP)

In the light of recent work on Maximum Non-extensive Entropy Partitioning for facial intensities[8], we further explore the use of non-extensive entropy[4] in the Maximum Entropy framework for foreground object detection.

We first describe below the maximum entropy partitioning for automated image thresholding[5–7] to divide the image into a bright segment, and a dark segment based on the uniform distribution of pixel intensity values within the two partitions.

Consider an image with L grey levels having a probability distribution $p_1, p_2, ...,p_L$ which is segregated into that belonging to the dark foreground object $p_1, p_2, ...,p_T$ and that to the brighter background $p_{+1}, p_{T+2}, ...,p_L$ based on an optimum threshold value T. The computation of the threshold is the ultimate aim of the image thresholding problem. The maximum entropy principle states that the optimum threshold T is the one for which the sum of entropies of the two partitions of the histogram is maximum[7]. The probabilities are normalised within each partition prior to the entropy computation by dividing the total sum of probabilities within each partition.

$$H(S) = H\left(\frac{p_1, p_2, \ldots, p_T}{p_1 + p_2 + \ldots + p_T}\right) + H\left(\frac{p_{T+1}, p_{T+2}, \ldots, p_L}{p_{T+1} + p_{T+2} + \ldots + p_L}\right) \quad \ldots\ldots\ldots\ldots\ldots (1)$$
$$= H\left(p_1^N, p_2^N, \ldots, p_T^N\right) + H\left(p_{T+1}^N, p_{T+2}^N, ..., p_L^N\right)$$

Both the sets of normalised probabilities $\left\{\frac{p_1, p_2, \ldots, p_T}{p_1 + p_2 + \ldots + p_T}\right\}$ and $\left\{\frac{p_{T+1}, p_{T+2}, \ldots, p_L}{p_{T+1} + p_{T+2} + \ldots + p_L}\right\}$ in (1) sum to 1 and form complete probability distributions.

The non-extensive entropy with Gaussian gain proposed[4] is defined as

$$H(P) = \sum_{i=1}^{n} p_i e^{-p_i^2} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

Being exponential, it is non-linear, and the favourable characteristics of the Gaussian curve make it a good choice for representing the sparse but repeating spatial probability patterns in texture images[4]. The sum of the non-extensive entropy with Gaussian gain of the two partitions is given by

$$H(S) = \sum_{i=1}^{T} p_i^N e^{-\left(p_i^N\right)^2} + \sum_{i=T+1}^{L} p_i^N e^{-\left(p_i^N\right)^2} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

Let the probabilities in the first partition belonging to the dark foreground object emulate a uniform or equally likely distribution with probabilities. $p_1^N = p_2^N = \cdots = p_T^N = 1/T$ Let the background clutter also approximate a uniform distribution $p_{T+1}^N = p_{T+2}^N = \cdots = p_L^N = 1/(L-T-1)$.

The sum of entropies in (3) then reduces to

$$H(S) = \sum_{i=1}^{T} \frac{1}{T} e^{-(1/T)^2} + \sum_{i=T+1}^{L} \frac{1}{L-T-1} e^{-(1/(L-T-1))^2} \quad \ldots\ldots\ldots\ldots\ldots (4)$$
$$= e^{-(1/T)^2} + e^{-(1/(L-T-1))^2}$$

The above expression is maximised when $T \to 1$ or $(T \to L–1)$ resulting in a maximum value of

$$H(S) = e^{-(1)^2} + e^{-(1/L-1)^2} \quad \ldots\ldots\ldots\ldots\ldots (5).$$

Due to the bell-shaped non-linear nature of the Gaussian curve, it tends to highlight and identify the low probability terms lying in the bell of the Gaussian [second term in (5)] and subdue the high probability events that lie along the

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.107 on dated 4-Jun-2018



(a) (b) (c) (d) (e)

Figure 1: The comparison of entropic-thresholding results for all methods; for the images labelled 1–6 in the order from top to bottom: (a) original image, (b) image thresholding using Shannon entropy, (c) image thresholding using Tsallis entropy, (d) image thresholding using Pal and Pal entropy, (e) image thresholding using the non-extensive entropy with Gaussian gain

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.107 on dated 4-Jun-2018

Table 1: Results of entropic image thresholding on various images in Figure 1 using various approaches *(Threshold values shown in Bold in the Table)*

Image	Image Thresholding Using Shannon Entropy Precision, Recall, FNo., Threshold	Image Thresholding Using Tsallis Entropy Precision, Recall, FNo., Threshold	Image Thresholding Using Pal and Pal Entropy Precision, Recall, FNo., Threshold	Image Thresholding Using the Non-Extensive Entropy with Gaussian GainPrecision, Recall, FNo., Threshold
1	0.78,0.57,0.66,**100**	1,0.35,0.51,**83**	0.78,0.57,0.66,**100**	0.78,0.57,0.66,**100**
2	1,0.53,0.69, **120**	1,0.51,0.67, **114**	1,0.19,0.31, **51**	1,0.18,0.30,**50**
3	0.61,0.76,0.67, **89**	0.41,0.82,0.55, **102**	0.63,0.75,0.68, **88**	0.63,0.75,0.68, **88**
4	0.56,0.96,0.70, **128**	0.56,0.96,0.70, **128**	0.56,0.95,0.70, **127**	0.59,0.95,0.72,**119**
5	0.45, 0.82, 0.58, **115**	0.41,0.84,0.55, **121**	0.51,0.79,0.61, **101**	0.62,0.75,0.68, **83**
6	0.63,0.97,0.76, **86**	0.57,0.99,0.72, **126**	0.63,0.97,0.76, **83**	0.65,0.97,0.77, **78**

bell of the Gaussian curve [first term in (5)]. Thus, the equiprobable darker pixels ($T \rightarrow 1$) pertaining to the foreground object are segmented out, as they fall in the bell of the Gaussian curve and emulate a uniform probability distribution. The remnant brighter image intensities constituting the background clutter fall along the bell of the Gaussian, and their relative probabilities are insignificant to cause a change in the maximum entropy position in (4). The six example images analysed in this paper are from the benchmark PASCAL VOC 2012 challenge dataset and consists of a varied assortment of dark objects against a bright background. We compare our results with that of the existing entropy-based image thresholding techniques as that of Wong and Sahoo's work based on Shannon entropy[7], Albuquerque *et al.*'s work based onTsallisentropy[6] and Pal and Pal entropy[5]. The programming is done on MATLAB 7.8 platform in a 2.6-GHz PC. The original images are labelled images 1 to 6, and the segmented results using all methods are shown in Figure 1. In all the cases, the darkest shades associated with the foreground object are segmented out more or less accurately by the non-extensive entropy. The high precision-recall values and the *F* measure for all the results summarised in Table 1 reiterate that the segmentation of dark foreground objects by the proposed method is in conformity to the ground truth annotations, and our results outperform that of all other entropy-thresholding methods in this regard.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.107 on dated 4-Jun-2018

3. CONCLUSION

We define and implement the maximum non-extensive entropy principle for segmenting the darkest shades associated with the foreground object. The sum of the two partitions will give a positive spike when the darkest shades in the image emulate a uniform probability distribution. Higher segmentation accuracies of segmentation of the dark objects against the brighter backgrounds in the image prove the efficiency of our approach.

REFERENCES

[1] Zhu H, Meng F, Cai J, Lu S. Beyond pixels: a comprehensive survey from bottom-up to semantic image segmentation and cosegmentation. Journal of Visual Communication and Image Representation 2016;34:12–27.

[2] Tian Y, Li J, Yu S, Huang T. Learning complementary saliency priors for foreground object segmentation in complex scenes. International Journal of Computer Vision 2015;111(2):153–70.

[3] Chang C-I, Du Y, Wang J, Guo S-M, Thouin PD. Survey and comparative analysis of entropy and relative entropy thresholding techniques. IEE Proceedings Vision, Image and Signal Processing 2006;153(6):837–50.

[4] Susan S, Hanmandlu M. A non-extensive entropy feature and its application to texture classification. Neurocomputing 2013;120:214–25.

[5] Pal N, Pal S. Entropy– anew definition and its applications. IEEE Transactions on Systems, Man, and Cybernetics1991;21(Sept (5)): 1260-1270.

[6] Portes de Albuquerque M, Esquef IA, Gesualdi Mello AR, Portes de Albuquerque M. Image thresholding using Tsallis entropy. Pattern Recognition Letters 2004;25(9):1059–65.

[7] Wong AKC, Sahoo PK. A gray-level threshold selection method based on maximum entropy principle. IEEE Transactions on Systems, Man, and Cybernetics 1989;19(4):866–71.

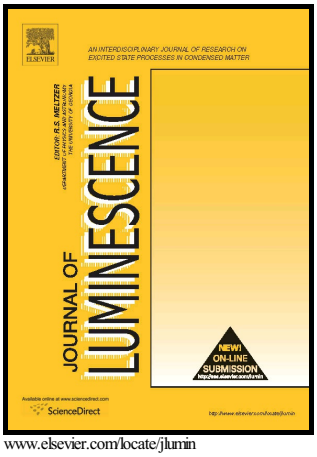
[8] Susan S, Aggarwal N, Chand S, Gupta A. Image coding based on maximum entropy partitioning for identifying improbable intensities related to facial expressions. Sadhana 2016;41(12):1393–406.

www.IndianJournals.com
Members Copy, Not for Commercial Sale
Downloaded From IP - 14.139.251.107 on dated 4-Jun-2018

Author's Accepted Manuscript

Sensitization of Nd^{3+} by 4f-5d transition of Ce^{3+} in $Ba_2Y(BO_3)_2Cl$ Phosphor for the Prospective NIR Applications

R.A. Talewar, Sk. Mahamuda, A.S. Rao, C.P. Joshi, S.V. Moharil



www.elsevier.com/locate/jlumin

PII: S0022-2313(18)30402-2
DOI: https://doi.org/10.1016/j.jlumin.2018.05.035
Reference: LUMIN15617

To appear in: *Journal of Luminescence*

Received date: 4 March 2018
Revised date: 11 May 2018
Accepted date: 11 May 2018

Cite this article as: R.A. Talewar, Sk. Mahamuda, A.S. Rao, C.P. Joshi and S.V. Moharil, Sensitization of Nd^{3+} by 4f-5d transition of Ce^{3+} in $Ba_2Y(BO_3)_2Cl$ Phosphor for the Prospective NIR Applications, *Journal of Luminescence,* https://doi.org/10.1016/j.jlumin.2018.05.035

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Sensitization of Nd^{3+} by 4f-5d transition of Ce^{3+} in $Ba_2Y(BO_3)_2Cl$ Phosphor for the Prospective NIR Applications

R.A. Talewar [1*], Sk. Mahamuda [1*], A.S. Rao [2], C.P. Joshi [3], S.V. Moharil [4]

[1] Department of Physics, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522 502, Guntur (Dt.), A.P, India.

[2] Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110 042, India.

[3] Physics Department, Shri Ramdeobaba College of Engineering and Management, Katol Road, Nagpur-440 013, India.

[4] Department of Physics, R.T.M. Nagpur University, Nagpur-440 010, India.

talewarrupesh@gmail.com (R.A. Talewar)

mahamuda.ss@gmail.com (Sk. Mahamuda)

* ***Corresponding author:*** Tel.: +91 9021092911

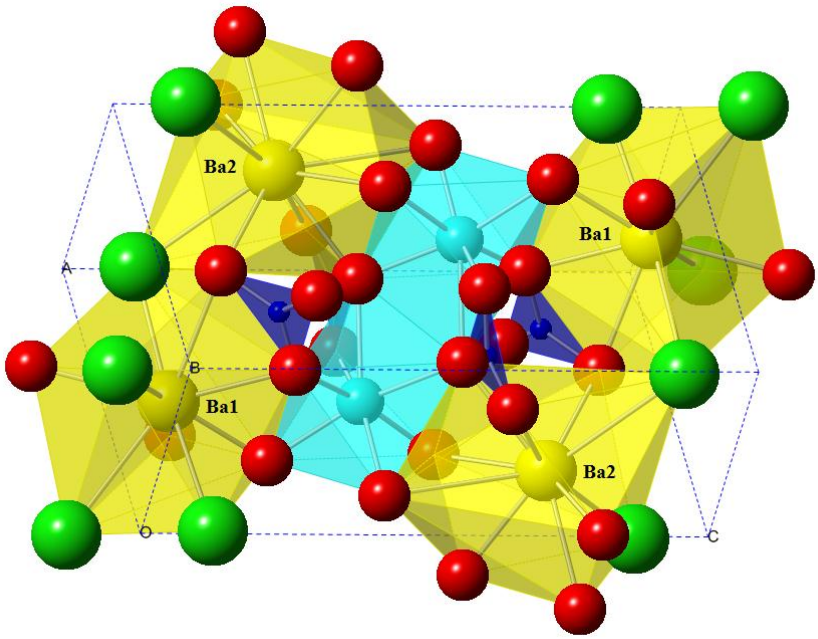
Accepted manuscript

Abstract

The $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphor was synthesized by conventional solid-state reaction method and characterized using XRD, SEM, photoluminescence (PL) excitation, PL emission and PL decay to understand the feasibility of using this phosphor for NIR emission. The XRD patterns recorded for the $Ba_2Y(BO_3)_2Cl$ phosphor reveal pure crystalline monoclinic phase belong to a space group $P2_1/m$. The surface morphology of the as prepared phosphor has been investigated using SEM. When excited with UV, the phosphor gives broadband emission at 420 nm, which corresponds to the allowed 5d→4f transition of Ce^{3+} ions and intense NIR emissions in the range 800-1400 nm, which are assigned to the characteristic $^4F_{3/2}\rightarrow{}^4I_{9/2,11/2,13/2}$ transitions of Nd^{3+} ions. The dependence of visible and NIR emissions, decay lifetime and energy transfer efficiency (η_{ETE}) were investigated in detail. Excitation, emission spectra and decay curves have

been measured to prove the energy transfer from Ce^{3+} to Nd^{3+}. The fluorescence decay measurements and energy transfer efficiencies have been estimated and the mechanisms for the energy transfer between Ce^{3+} and Nd^{3+} have been proposed. The results indicate that sensitization of Nd^{3+} is possible via 4f-5d transition of Ce^{3+}.

Graphical abstract



Coordination of cations in $Ba_2Y(BO_3)_2Cl$ phosphor

ACCEPTED MANUSCRIPT

Keywords: Near-infrared; Energy transfer; Ce^{3+}-Nd^{3+}; Sensitization; Phosphors; Photoluminescence

Accepted manuscript

ACCEPTED MANUSCRIPT

1. Introduction

Nowadays, near-infrared (NIR) emitting luminescent materials are of fundamental and technological importance because of a broad diversity of high-value applications in light emitting diodes, optical communications, optical sensors, lasers, photovoltaic and bio-imaging [1-6]. In luminescent materials, rare earth (RE) ions play an important role due to their abundant energy levels with emission in different regions based on their 4f-4f or 5d-4f transitions. RE ions like Eu^{3+}, Tb^{3+}, Sm^{3+}, Pr^{3+}, Ho^{3+} are popular for their emission in visible region and therefore are useful for applications like fluorescent lamps [7], plasma display panels [8] and white light emitting diodes [9]. There are few interesting reports available for NIR emission of RE ions (such as Pr^{3+} and Ho^{3+}) doped hosts like Sr_2SiO_4 [10], $CaLaGa_3S_6O$ [11] and YVO_4 [12]. Among the RE ions, Er^{3+}, Yb^{3+}, Nd^{3+} and Tm^{3+} appear to be important candidates for NIR emission due to their favourable energy level schemes [13]. These RE ions doped various host materials have been extensively studied and used to generate NIR emissions for various applications [14-16]. Despite the great potential for applications, the development of efficient NIR emitters has long been hampered by their much lower emission efficiency compared with that of visible emitters.

In quite recent years, the interest in developing the suitable materials for better NIR emission around 1μm has been escalating because of its applications in diversified fields such as lasers, optical fiber amplifiers, photovoltaics and bio-imaging [17-20]. Among various RE ions capable of emitting in the NIR region, Nd^{3+} is particularly noteworthy since its emission locates at around 1060 nm [21]. Trivalent neodymium has unique properties such as numerous electronic excited states suitable for optical pumping in the ultraviolet to the near-infrared wavelength range and strong NIR luminescence. As a result, Nd^{3+} doped materials have been extensively studied as effective NIR emitters with potential applications in lasers and optical temperature sensors [22,23]. Transitions of Nd^{3+} ion involve the $4f^N$ intra band transitions from $^4F_{3/2}$ state to ^4I_J (J = 9/2, 11/2, 13/2) states [24]. Crystalline and non-crystalline materials doped with Nd^{3+} either single or co-doped with another suitable RE ions are capable of giving emission in NIR region [25-28]. The single RE ion NIR emission is not significant for practical application because the transitions between 4f levels of trivalent lanthanide ions are forbidden. The oscillator strengths of 4f–4f transitions are typically of the order of 10^{-6}; dipole allowed transitions, by comparison, may have oscillator strength of up to unity [29,30]. Thus, the absorption strength of lanthanide 4f-4f transitions is very weak [19]. NIR emission of Nd^{3+} can also be sensitized using RE couples like Nd^{3+}-Yb^{3+} [31], Er^{3+}-Nd^{3+} [32] in various matrices. However, the weak

Accepted manuscript

absorption cross-section of the 4f-4f transitions of these sensitizers makes them less attractive for the above said applications. An effective way to tackle this inherent problem of f-f transitions is to use sensitizers like Ce^{3+}, Eu^{2+} or Bi^{3+} with allowed transitions. Over the years, Ce^{3+} and Eu^{2+} ions have been widely investigated as efficient activators in phosphors for n-UV and blue LEDs because they usually have a strong excitation band covering the emissions from n-UV and blue regions. These broad excitation bands are due to the allowed inter-configurational 4f–5d transitions with high oscillator strength [33, 34]. Also, the 5d→4f emissions of Ce^{3+} or Eu^{2+} can vary from long-wavelength ultra-violet to red light due to their strong dependence on the host composition, crystal structure, the lattice symmetry and so on [35,36]. Bi^{3+} may also serve as sensitizer for Nd^{3+} with the transitions involving $^1S_0 \rightarrow {}^3P_1$ for intense and broadband absorption [37]. Therefore, if the Ce^{3+}/ Eu^{2+} (with strong 4f–5d broad absorption) or Bi^{3+} ($^1S_0 \rightarrow {}^3P_1$ allowed transition) and the Nd^{3+} (4f–4f sharp line emission) are combined to form a sensitizer–activator pair, the co-doped phosphors would possibly overcome all the aforementioned drawbacks.

With the aim of preparing NIR emitting phosphor and sensitization from 4f-5d allowed transition, we chose Ce^{3+} as a sensitizer for Nd^{3+}. Over the years, Ce^{3+} ion has been widely investigated as efficient activator in phosphors for n-UV and blue LEDs because they usually have a strong excitation band covering the emission from n-UV and blue LEDs, and an intense emission band, due to the allowed inter-configurational 4f-5d transitions with high oscillator strength [38]. Also, the 5d→4f emissions of Ce^{3+} can vary from long-wavelength ultra-violet to yellow due to their strong dependence on the host composition, crystal structure, the lattice symmetry and so on [39]. For this we prepared $Ba_2Y(BO_3)_2Cl$ co-doped with Ce^{3+}-Nd^{3+} through two step solid-state reaction. According to the results published by Jing et. al, the compound $Ba_2Y(BO_3)_2Cl$ offers excellent chemical and physical stability. The emission of Ce^{3+} in $Ba_2Y(BO_3)_2Cl$ is centred at about 415 nm with a broad strong absorption band ranging from 300 to 400 nm [40]. It was also found suitable for studying energy transfers, e.g. Bi^{3+}-Yb^{3+} [41], Yb^{3+}-Er^{3+} [42], Bi^{3+}-Eu^{3+} [43], Ce^{3+}-Tb^{3+} [44], Ce^{3+}-Yb^{3+} [45], etc. $Ba_2Ln(BO_3)_2Cl$ (Ln= Ho, Yb) type compound was first reported by Khamaganova et. al in 1989 [46]. Later Schipper et. al discussed the luminescence of Ce^{3+}, Eu^{2+}, Eu^{3+} and Gd^{3+} in the compound $Ba_2Ln(BO_3)_2Cl$ (Ln=Gd, Y, Lu) [47]. Thereafter, number of reports appeared focussing on rare earth doped $Ba_2Y(BO_3)_2Cl$. Jing et. al investigated effects of Eu^{3+} sites on photoluminescence in $Ba_2Ln(BO_3)_2Cl$ (Ln = Gd, Y) hosts [48]. Recently, number of papers appeared on the luminescence of Bi^{3+}, Eu^{3+} and Eu^{2+} doped in $Ba_2Ln(BO_3)_2Cl$ phosphor [49-51]. The emphasis

Accepted manuscript

ACCEPTED MANUSCRIPT

of these works was on visible region and there were no reports on NIR emission corresponding to energy transfer from Ce^{3+} to Nd^{3+} in $Ba_2Y(BO_3)_2Cl$ phosphor.

In the present work we are aiming to prepare an efficient phosphor doped with an activator and sensitizer for NIR emission. We are choosing $Ba_2Y(BO_3)_2Cl$ phosphor for this purpose as it is capable of converting UV to near infrared. We are choosing Ce^{3+} as a sensitizer and Nd^{3+} as an activator ion. We prepared a $Ba_2Y(BO_3)_2Cl$ phosphor co-doped with Ce^{3+}-Nd^{3+} through a two-step solid-state reaction process. In this paper, we explained clearly the energy transfer mechanism that took place between sensitizer (Ce^{3+}) and activator (Nd^{3+}) in the co-doped $Ba_2Y(BO_3)_2Cl$ phosphor. To the best of our knowledge, energy transfer from Ce^{3+}–Nd^{3+} co-doped $Ba_2Y(BO_3)_2Cl$ phosphor has not yet been reported in literature.

2. Experimental

$Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ powder samples with different doping concentrations were prepared by conventional solid state reaction technique. All the starting materials viz. $BaCO_3$, $BaCl_2$, Y_2O_3, $Ce_2(CO_3)_3$, $Nd_2(SO_4)_3.8H_2O$ and H_3BO_3 were of Analytical Reagent grade. Excess 5% H_3BO_3 above the stoichiometric proportions was used. The powders were thoroughly mixed in an agate mortar by grinding and the mixture was transferred into crucibles, pre-calcined at 600°C for 3 hours in air and finally sintered at 950°C for 3 hours under reducing atmosphere, followed by an additional grinding.

The crystalline phases of the synthesized samples were identified by x-ray diffraction (XRD) patterns recorded on XPERT-PRO diffractometer using Cu K_α (λ=1.54059 Å) radiation. The surface morphology of the sample was studied by using scanning electron microscope (Zeiss EVO 18). The UV-vis-NIR PL and PL excitation (PLE) spectra were measured by a Hitachi-F7000 florescence spectrophotometer and Photon Technology International QM-51 NIR spectrophotometer. PL decay curves were measured by using a time-correlated single photon counting (TCSPC) instrument from IBH Horiba-JY (United Kingdom) with 374 nm LED as an excitation source. All the measurements were carried out at room temperature.

3. Results and discussion

Fig. 1 shows the XRD patterns of $Ba_2Y(BO_3)_2Cl:Ce^{3+}$ sintered at 950°C for 3 hours under reducing atmosphere. In previous works the patterns were compared with those for $Ba_2Yb(BO_3)_2Cl$. Atomic coordinates for $Ba_2Y(BO_3)_2Cl$ are known [51]. We computed XRD pattern from this information using Vesta program [52]. The recorded XRD patterns are

Accepted manuscript

compared with the computed one. The XRD patterns of the samples reveal that all the peaks of the sample can be identified with the pure crystalline monoclinic $Ba_2Y(BO_3)_2Cl$ belonging to space group $P2_1/m$, and the lattice parameter values are a = 6.424 Å, b = 5.307 Å, c = 11.279 Å. The result clearly indicates that the obtained samples are single-phase and acquires the same structure as $Ba_2Y(BO_3)_2Cl$. The doping of Ce^{3+} does not cause any significant change to the host structure. Fig. 2 shows the crystal structure of $Ba_2Y(BO_3)_2Cl$, emphasizing the coordination environment of Ba^{2+}, Y^{3+} and B^{3+} atom. It can be observed that the structural unit of $Ba_2Y(BO_3)_2Cl$ contains isolated triangular BO_3 groups and (Ba–Cl, O) and Y–O polyhedra. The Ba^{2+} atoms occupy two sites, Ba(1) is 9-fold coordinated by 5O and 4Cl [$Ba(1)O_5Cl_4$], while Ba(2) is 10-fold coordinated with 8O and 2Cl [$Ba(2)O_8Cl_2$]. Y^{3+} is 7-coordinated by oxygen with site symmetry C_S [46]. In view of the effective ionic radii of cations Ba^{2+}, 1.47 Å [$Ba(1)O_5Cl_4$] and 1.52 Å [$Ba(2)O_8Cl_2$]; Y^{3+}, 0.96 Å; Ce^{3+}, 1.07 Å Ce^{3+} can occupy the Ba^{2+} or Y^{3+} substitutional sites [53]. In order to analyze the surface morphology of sample, we have pursued scanning electron microscopic (SEM) study. Fig. 3 (a) depicts the surface micrograph of the sample, micron size grains can be easily visualized with agglomeration. To estimate the size distribution, we selected the region where there is a less agglomeration of particles (indicated by Yellow Square). Fig. 3 (b) indicates the magnified view of the yellow square region of SEM image. From magnified image, we measured the size of each particle and plotted in terms of size distribution [see Inset of Fig. 3 (b)]. Size distribution of particles yields mean particle size in the range of 1.4-1.5 μm [Inset of Fig. 3 (b)]. Most of the commercially available phosphors are also exhibiting particle size of the order of few micrometres [54].

The PLE and PL spectra of Ce^{3+} doped $Ba_2Y(BO_3)_2Cl$ with Ce^{3+} concentration varying from 1 mol% to 4 mol% are shown in Fig. 4. Corresponding to the 420 nm emission, a broad excitation band centered at 360 nm with a shoulder at 340 nm can be observed. These excitation bands correspond to the transitions from 4f level ($^2F_{5/2}$) to different crystal field-split levels of the $5d^1$ states for the Ce^{3+} ions [55]. The PL spectrum under 360 nm excitation shows a typical asymmetric double humped emission with maxima around 385 nm and 420 nm. These emissions are in good agreement with the reported value for $Ba_2Y(BO_3)_2Cl:Ce^{3+}$ [40]. The typical double-band emission can be assigned to the parity allowed transitions from the lowest component of 5d state to $^2F_{5/2}$ and $^2F_{7/2}$ levels of Ce^{3+} ions due to the strong coupling of the 5d electron with host lattice [39]. At low Ce^{3+} doping concentrations, these two bands are observed whereas for the increased doping concentration, the doublet structure is not clearly resolved due to the

Accepted manuscript

inhomogeneous broadening of the Ce^{3+} emission band resulting from the multiple random distribution site occupancy. Fig. 5 shows the emission spectra of the Ce^{3+} doped $Ba_2Y(BO_3)_2Cl$ phosphor decomposed by Gaussian deconvolution into two Gaussian profiles with peaks centered at 388 nm and 418 nm (corresponding to 25773 cm^{-1} and 23923 cm^{-1}, respectively), which can be attributed to the transition from the lowest 5d level to the $^2F_{5/2}$ and $^2F_{7/2}$ levels of Ce^{3+}. The energy difference of two sub-bands is approximately 1849 cm^{-1}, which is close to the ^2F_J (J = 7/2, 5/2) energy gap of Ce^{3+} (≈2000 cm^{-1}) in most Ce^{3+} activated phosphors [56,57]. The effect of doping concentration of Ce^{3+} (1-4 mol%) on the emission intensity of $Ba_2Y(BO_3)_2Cl$ phosphors was also investigated. The optimum concentration of Ce^{3+} in $Ba_2Y(BO_3)_2Cl$ was found to be 2 mol%. Inset in Fig. 4 depicts the concentration quenching effect. This concentration quenching may be due to energy transfer (ET) process among Ce^{3+} ions, since it occupies three different sites in the host lattice [51]. By carefully observing the emission spectra, it can be seen that at higher Ce^{3+} doping concentration, the double band emission due to the transitions of the two $^2F_{5/2}$ and $^2F_{7/2}$ ground state components cannot be resolved which may be due to the spectral overlap or the inhomogeneous broadening of different Ce^{3+} sites in the lattice [58]. Similar type of observation is reported by Jing et al. and Wang et al. [40,59].

The PLE and PL spectra of Ce^{3+}-Nd^{3+} co-doped $Ba_2Y(BO_3)_2Cl$ are shown in Fig. 6. In $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphors, the concentration of Ce^{3+} was fixed at 2 mol% and the concentrations of co-dopant Nd^{3+} was varied from 1-12 mol%. Under 362 nm excitation, the broadband emission at 420 nm corresponding to the 5d→4f transitions of Ce^{3+} and intense NIR emissions in the range 800-1400 nm, which are assigned to the transitions from $^4I_{9/2,11/2,13/2}$ levels of Nd^{3+} ions, are observed. The excitation spectrum from 300 to 400 nm by monitoring the NIR emission at 1060 nm is almost the same as that monitored at 420 nm emission. This is also consistent with the excitation spectrum for $Ba_2Y(BO_3)_2Cl:Ce^{3+}$ sample, which is assigned to the 4f→5d transition of the Ce^{3+} ions. In addition to the 4f→5d band of Ce^{3+}, various lines due to f-f transitions of Nd^{3+} in the range 400-800 nm were also observed in the excitation spectrum. These absorption lines of Nd^{3+} were at 427, 436, 464, 476, 515, 534, 579, 591, 619, 686, 745 and 765 nm, which are attributed to the $^4I_{9/2}\rightarrow{}^2D_{5/2}$ (427 nm), $^4I_{9/2}\rightarrow{}^2P_{1/2}$ (436 nm), $^4I_{9/2}\rightarrow{}^4G_{11/2}$ (464 nm), $^4I_{9/2}\rightarrow{}^4G_{9/2}$ (476 nm), $^4I_{9/2}\rightarrow{}^2G_{9/2}$ (515 nm), $^4I_{9/2}\rightarrow{}^4G_{7/2}$ (534 nm), $^4I_{9/2}\rightarrow{}^2G_{7/2}$ (579 nm), $^4I_{9/2}\rightarrow{}^4F_{5/2}$ (591 nm), $^4I_{9/2}\rightarrow{}^2H_{11/2}$ (619 nm), $^4I_{9/2}\rightarrow{}^4F_{9/2}$ (686 nm), $^4I_{9/2}\rightarrow{}^4S_{3/2}$ (745 nm) and $^4I_{9/2}\rightarrow{}^4F_{7/2}$ (756 nm) transitions of Nd^{3+}, respectively [60,61]. It can be noticed that the emission spectrum of $Ba_2Y(BO_3)_2Cl:Ce^{3+}$ overlapped well with four sharp excitation peaks of Nd^{3+} singly

ACCEPTED MANUSCRIPT

doped $Ba_2Y(BO_3)_2Cl$ at 387 nm, 427 nm, 436 nm and 467 nm, indicating the possible ET from Ce^{3+} to Nd^{3+} ions. Excitation and emission spectra of singly doped $Ba_2Y(BO_3)_2Cl:Nd^{3+}$ phosphors are shown in Fig. 7. Intense characteristic emission lines of Nd^{3+} at 890 nm, 1060 nm and 1337 nm which are assigned to the $^4F_{3/2} \rightarrow {}^4I_j$ (j = 9/2, 11/2 and 13/2) transitions of Nd^{3+} were observed under 823 nm excitation. Excitation spectrum consists of several lines corresponding to various f-f transitions of Nd^{3+}. Fig. 7 shows that for singly doped $Ba_2Y(BO_3)_2Cl:Nd^{3+}$, strong excitation exists around 590 and 748 nm in form of relatively narrow lines. For many optoelectronic applications, it is desirable to have a broad band excitation down to the nUV region of the electromagnetic spectrum. Especially such nUV broad excitation is quite useful in photovoltaic solar cells in conserving UV radiation of solar spectrum more efficiently [19]. However prolonged exposure of nUV radiation will be detrimental for the longevity of solar cells. Hence these radiations should be effectively removed for increasing the life span of solar cells. For this it is necessary to use dopants like Eu^{2+} or Ce^{3+} having strong and broad excitation bands to sensitize Nd^{3+} emission [62,63]. From Fig. 6 it is conspicuous that in the doubly doped $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphor, the broad excitation is extended down to nUV region. Thus $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphor can be effectively used to convert nUV radiations of the solar spectrum in to NIR and is more useful for solar cell applications. The $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphor can also be used as an efficient NIR emitter when excited with nUV LED's. One can also obtain compact, diode pumped NIR sources.

To establish the existence of ET from Ce^{3+} to Nd^{3+}, the dependence of 5d→4f emission of Ce^{3+} and NIR emission of Nd^{3+} on the Nd^{3+} co-doping concentration in $Ba_{2-0.04}Y_{1-x}(BO_3)_2Cl:0.04Ce^{3+}$, xNd^{3+} (x = 0.01, 0.03, 0.06, 0.09 and 0.12) are shown in Fig. 8. It can be seen clearly that with increase in Nd^{3+} ions concentration from 1 mol% to 12 mol%, the PL intensity of Ce^{3+} emission decreases continuously, while the emission intensity of Nd^{3+} first increases rapidly with x increasing up to 0.06 and then decreases with further increase due to the concentration quenching effect. The above phenomenon offers direct evidence of ET from Ce^{3+} to Nd^{3+}.

The decay process of $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphors was also investigated to confirm the energy transfer from $Ce^{3+} \rightarrow Nd^{3+}$. The luminescence decay curves of Ce^{3+} (excited at 374 nm and monitored at 420 nm) of $Ba_2Y(BO_3)_2Cl:0.04Ce^{3+}$, xNd^{3+} (x = 1, 3, 6 and 9 mol%) phosphors are shown in Fig. 9. The corresponding luminescence decay curves were well fitted with a second-order exponential decay mode using the following equation [64,65]

ACCEPTED MANUSCRIPT

$I = a_1 \exp^{(\frac{-t}{\tau 1})} + a_2 \exp^{(\frac{-t}{\tau 2})}$

where, I is the luminescence intensity at time t, a_1 and a_2 are the amplitudes of decay constants and τ_1 and τ_2 are the luminescence lifetimes for fast and slow channels of decay respectively. The values of a_1, τ_1, a_2, and τ_2 are obtained as shown in Table 1. The average lifetime τ_{avg} in case of a second-order exponential decay were calculated using the expression given below and the values are given in Table 1.

$\tau_{avg} = \frac{a_1\tau_1^2 + a_2\tau_2^2}{a_1\tau_1 + a_2\tau_2}$

From the Table 1, it is observed that the average lifetime of Ce^{3+} doped $Ba_2Y(BO_3)_2Cl$ phosphor was found to be 25.37 ns. The lifetime of the order of nanosecond is one of the characteristics of the electric-dipole allowed 5d–4f transition of Ce^{3+} [66,67]. However, it is observed that with increase of Nd^{3+} ion concentration the lifetime of Ce^{3+} decreases monotonously, implying introduction of extra decay pathways. This decrease in the lifetime of Ce^{3+} in $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ phosphor with the increase of Nd^{3+} concentration indicates that there is an energy transfer from Ce^{3+} to Nd^{3+} ions. The energy transfer efficiency (η_{ETE}) from Ce^{3+} to Nd^{3+} in the system $Ba_2Y(BO_3)_2Cl$ can be calculated by using the following equation [68,69],

$$\eta_{ETE} = 1 - \frac{\tau_x}{\tau_0} \quad (1)$$

Where τ_0 is the lifetime of the sensitizer Ce^{3+} in the sample without Nd^{3+} co-doping, and τ_x is the lifetime of Ce^{3+} along with Nd^{3+} co-doping. The obtained average lifetimes with Ce^{3+} singly doped and Ce^{3+}, Nd^{3+} co doped $Ba_2Y(BO_3)_2Cl$ with energy transfer efficiency are presented in Table 1.

Fig. 10 represents the schematic energy level diagram of $Ba_2Y(BO_3)_2Cl$: Ce^{3+}, Nd^{3+} phosphors, and possible energy transfer mechanism. After excitation to the 5d state of Ce^{3+}, the excited levels can either relax radiatively to the 4f ground state of Ce^{3+}, producing broadband luminescence in the visible region or transfer of energy to the acceptor levels of Nd^{3+}. The energy transfer process may follow the following process:

Ce^{3+}: 5d→4f to Nd^{3+}: $^2P_{3/2}$, $^2D_{5/2}$ and $^2P_{1/2}$.

Then, the electrons can relax through non-radiative decay to $^4F_{5/2}$ or $^4F_{3/2}$ levels, and then decay radiatively to the ^4I_j (j = 9/2, 11/2 and 13/2) levels of Nd^{3+}, emitting at 890 nm, 1060 nm and 1337 nm, respectively.

4. Conclusions

Accepted manuscript

ACCEPTED MANUSCRIPT

In summary, NIR emitting phosphors, $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ were synthesized by solid state reaction and investigated for the possible energy transfer. The XRD pattern recorded for $Ba_2Y(BO_3)_2Cl$ phosphor reveals monoclinic phase of $P2_1/m$ space group. The SEM micrograph reveals inhomogeneous, irregular morphology with some agglomeration in the sample. The PLE and PL spectra and decay lifetime measurements evidently proves that there is an energy transfer from Ce^{3+} to Nd^{3+} ions with the energy transfer efficiency (η_{ETE}) reaching up to 84.8%. We propose that the ET from Ce^{3+} to Nd^{3+} in $Ba_2Y(BO_3)_2Cl$ occurs mainly due to first order energy transfer mechanism i.e. cross relaxation involving phonon-assisted processes. An efficient energy transfer from Ce^{3+} to Nd^{3+} leads to give intense NIR emission. The strongest NIR emission of Nd^{3+} is observed for concentrations of 2 mol% for Ce^{3+} and 6 mol% for Nd^{3+}. The results indicate that sensitization of Nd^{3+} is possible via 4f-5d transition of Ce^{3+}. Thus this phosphor may serve for Nd^{3+}-based optical and photonic devices.

Acknowledgement

This work was financially supported by Department of Science and Technology, Govt. of India (SR/S2/CMP-008/2012 dated 20/05/2013 and ECR/2016/000376). The authors would like to acknowledge Department of Physics, RTM Nagpur University, Nagpur for providing SEM facility. The authors would like to acknowledge Dr. Anindya Datta, Professor at Department of Chemistry, IIT-Bombay, India and Prof. Sujit Kumar Ghosh, Visvesvaraya National Institute of Technology (VNIT), Nagpur, India for their valuable discussions and for providing Fluorescence Decay Measurements facilities.

Accepted manuscript

ACCEPTED MANUSCRIPT

References

[1] Kiet Tuong Ly, Ren-Wu Chen-Cheng, Hao-Wu Lin, Yu-Jeng Shiau, Shih-Hung Liu, Pi-Tai Chou, Cheng-Si Tsao, Yu-Ching Huang and Yun Chi, Nature Photonics 11 (2017) 63–68

[2] Chenglai Jia, Bhavin J. Shastri, Nurmemet Abdukerim, Martin Rochette, Paul R. Prucnal, Mohammed Saad and Lawrence R. Chen, Sci. Rep. 6 (2016) 36071

[3] Legang Li, Junyi Bian, Qing Jiao, Zijun Liu, Shixun Dai and Changgui Lin, Sci. Rep. 6 (2016) 37577

[4] Xiangfu Wang, Ye Wang, Jose Marques-Hueso and Xiaohong Yan, Sci. Rep. 7 (2017) 758

[5] Zijun Liu, Jinyan Li, Luyun Yang, Qiaoqiao Chen, Yingbo Chu and Nengli Dai, Sol. Energy Mater. Sol. Cells 122 (2014) 46–50

[6] Abdukader Abdukayum, Jia-Tong Chen, Qiang Zhao and Xiu-Ping Yan, J. Am. Chem. Soc. 135 (2013) 14125–14133

[7] T. Jüstel, Hans Nikol and Cees Ronda, Angew. Chem. Int. Ed. 37 (1998) 3084–3103

[8] Chang-Hong Kim, Il-Eok Kwon, Cheol-Hee Park, Young-Ju Hwang, Hyun-Sook Bae, Byung-Yong Yu, Chong-Hong Pyun and Guang-Yan Hong, J. Alloys Compd. 311 (2000) 33–39

[9] P.J. Yadav, C.P. Joshi and S.V. Moharil, J. Lumin. 136 (2013) 1–4

[10] Ye Li, Qiuling Yu, Lin Huang, Jing Wang and Qiang Su, Opt. Mater. Express 4 (2014) 227–233

[11] G. G. Zhang, C. M. Liu, J. Wang, X. J. Kuang and Q. Su, J. Mater. Chem. 22 (2012) 2226–2232

[12] D.C. Yu, X.Y. Huang, S. Ye, M.Y. Peng and Q.Y. Zhang, Sol. Energy Mater. Sol. Cells 101 (2012) 303–307

[13] G. H. Dieke and H. M. Crosswhite, Appl. Opt. 2 (1963) 675–686

[14] G.S. Li, C.M.Zhang, P.F.Zhu, C.Jiang, P.Song and K.Zhu, Ceram. Int. 42 (2016) 5558–5561

[15] D. Serrano, A. Braud, J.-L. Doualan, P. Camy, A. Benayad, V. Ménard and R. Moncorgé, Opt. Mater. 33 (2011) 1028–1031

[16] U.R. Rodríguez-Mendoza and F. Lahoz, J. Lumin. 179 (2016) 40–43

[17] F. Rivera-López, P. Babu, Ch. Basavapoornima, C.K. Jayasankar and V. Lavín, J. Appl. Phys. 109 (2011) 123514

[18] Yehui Liu, Yibo Wang, Lei Liao, Xing Luo, Xiongwei Hu, Fangfang Zhang, Haiqing Li, Jinggang Peng, Lvyun Yang and Jinyan Li, Laser Phys. Lett. 14 (2017) 045102

Accepted manuscript

ACCEPTED MANUSCRIPT

[19] Bryan M. van der Ende, Linda Aarts and Andries Meijerink, Phys. Chem. Chem. Phys. 11 (2009) 11081–11095

[20] M. Pokhrel, L.C. Mimun, B. Yust, G.A. Kumar, A. Dhanale, L. Tang and D.K. Sardar, Nanoscale. 6 (2014) 1667–1674

[21] C. Nageswara Raju, C. Adinarayana Reddy, S. Sailaja, Hyo Jin Seo and B. Sudhakar Reddy, J. Mater. Sci. 47 (2012) 772–778

[22] R.E. Avalyan, G.S. Gevorkyan, A.A. Glushko and K.B. Petrosyan, Infrared Phys. 29 (1989) 213–214

[23] Nikifor Rakov and Glauco S. Maciel, J. Appl. Phys. 121 (2017) 113103

[24] A.A. Pathak, R.A. Talewar, C.P. Joshi and S.V. Moharil, J. Lumin. 179 (2016) 350–354

[25] John B. Gruber, Dhiraj K. Sardar, Raylon M. Yow, Toomas H. Allik and Bahram Zandi, J. Appl. Phys. 96 (2004) 3050–3056

[26] Jair Ricardo de Moraes, Fernando Rodrigues da Silva, Laercio Gomes, Vera Lucia Mazzocchi, Carlos Benedicto Ramos Parente and Sonia Licia Baldochi, CrystEngComm 15 (2013) 2260–2268

[27] Jian Xin Meng, Jin Qing Li, Zhao Pu Shi, and Kok Wai Cheah, Appl. Phys. Lett. 93 (2008) 221908

[28] B. Viana, M. Palazzi and O. LeFol, J. Non-Cryst. Solids 215 (1997) 96–102

[29] J. Rubio O, J. Phys. Chem. Solids 52 (1991) 101–174

[30] Oscar L. Malta and Luís D. Carlos, Quim. Nova 26 (2003) 889–895

[31] Rolindes Balda, Jose Ignacio Peña, M. Angeles Arriandiaga and Joaquín Fernández, Opt. Express 18 (2010) 13842–13850

[32] Shen Xiang, Nie Qiuhua , Xu Tiefeng, Dai Shixun and Wang Xunsi, J. Rare Earths 26 (2008) 899–903

[33] Y.Q. Li, G. de With, H.T. Hintzen, J. Mater. Chem. 15 (2005) 4492–4496

[34] P. Dorenbos, J. Lumin. 104 (2003) 239–260

[35] M. Kottaisamy, P. Thiyagarajan, J. Mishra and M.S.R. Rao, Mater. Res. Bull. 43 (2008) 1657–1663

[36] A. Diaz and D.A. Keszler, Mater. Res. Bull. 31 (1996) 147–151

[37] X.Y. Huang, X.H. Ji, and Q.Y. Zhang, J. Am. Ceram. Soc. 94 (2011) 833–837

[38] P. Dorenbos, J. Phys.: Condens. Matter 15 (2003) 4797–4807

[39] P. Dorenbos, J. Lumin. 91 (2000) 155–176

Accepted manuscript

ACCEPTED MANUSCRIPT

[40] Heng Jing, Chongfeng Guo, Gongguo Zhang, Xiangying Su, Zheng Yang and Jung Hyun Jeong, J. Mater. Chem. 22 (2012) 13612–13618

[41] Anjun Huang, Zhengwen Yang, Chengye Yu, Zhuangzhuang Chai, Jianbei Qiu and Zhiguo Song, IEE Photonics Journal (2017) DOI: 10.1109/JPHOT.2017.2782842

[42] Anjun Huang, Zhengwen Yang, Chengye Yu, Jianbei Qiu and Zhiguo Song, J. Am. Ceram. Soc. 100 (2017) 4994–4998

[43] Anjun Huang, Zhengwen Yang, Chengye Yu, Zhuangzhuang Chai, Jianbei Qiu and Zhiguo Song, J. Phys. Chem. C 121 (2017) 5267–5276

[44] J. Sun, J. Lai, Z. Xia, X. Zhang, H. Liu and H. Du, Appl. Phys. B 107 (2012) 827–831

[45] Jin Zhao, Chongfeng Guo and and Ting Li, ECS J. Solid State Sci. Tech. 5 (2016) R3055–R3058

[46] T.N. Khamaganova, N.N. Nevskii and V.K. Trunov, Sov. Phys. Cryst. 34 (1989) 853 translation of Kristallografiya 34 (1989) 1422–1425

[47] W.J. Schipper and G. Blasse, J. Alloys Compd. 203 (1994) 267–269

[48] Heng Jing, Chongfeng Guo, Niumiao Zhang, Zhaoru Ren and Jintao Bai, ECS J. Solid State Sci. Tech. 2 (2013) R1–R4

[49] Anjun Huang, Zhengwen Yang, Chengye Yu, Zhuanzhuang Chai, Jianbei Qiu and Zhiguo Song, Mater. Lett. 185 (2016) 440–442

[50] Dongyan Yu, Yujun Liang, Zhanggen Xia, WenZhu Huang and Mengfei Zhang, J. Crystal Growth 397 (2014) 13–18

[51] Zhiguo Xia, Xiaoming Wang, Yingxia Wang, Libing Liao and Xiping Jing, Inorg. Chem. 50 (2011) 10134–10142

[52] K. Momma and F. Izumi J. Appl. Cryst. 44 (2011) 1272-1276

[53] R.D. Shannon, Acta Cryst. A32 (1976) 751–767

[54] RuijinYu, Yue Guo, Lili Wang, Hyeon Mi Noh, Byung Kee Moon, Byung Chun Choi and Jung Hyun Jeong, J. Lumin. 155 (2014) 317–321

[55] P. Dorenbos, Phys. Rev. B 65 (2002) 235110

[56] P. Dorenbos, J. Lumin. 91 (2000) 91–106

[57] D.K. Ingole, C.P. Joshi, S.V. Moharil, P.L. Muthal and S.M. Dhopte, J. Lumin. 131 (2011) 2499–2502

[58] J.J. Ju, T.Y. Kwon, S.I. Kim, J.H. Kim, M. Cha and S.I. Yun, Mater. Lett. 28 (1996) 149–153

Accepted manuscript

ACCEPTED MANUSCRIPT

[59] Tao Wang, Yihua Hu, Li Chen, Xiaojuan Wang and Guifang Ju, Appl. Phys. A 120 (2015) 301–308

[60] A.A. Pathak, R.A. Talewar, C.P. Joshi and S.V. Moharil, Opt. Mater. 64 (2017) 217–223

[61] A.S. Rao, Y.N. Ahammed, R.R. Reddy and T.V.R. Rao, Opt. Mater. 10 (1998) 245–252

[62] R.A. Talewar, C.P. Joshi and S.V. Moharil, J. Lumin. 197 (2018)1–6

[63] Jiajia Zhou, Yu Teng, Xiaofeng Liu, Zhijun Ma and Jianrong Qiu, J. Mater. Res. 26 (2011) 689–692

[64] C.H. Huang and T.M. Chen, Opt. Express 18 (2010) 5089–5099

[65] S. Buddhudu, M. Morita, S. Murakami and D. Rau, J. Lumin. 83–84 (1999) 199–203

[66] JinDeng Chen, Hai Guo, ZhengQuan Li, Hao Zhang and YiXi Zhuang, Opt. Mater. 32 (2010) 998–1001

[67] Xinguo Zhang, Liya Zhou, Qi Pang, Jianxin Shi and Menglian Gong, J. Phys. Chem. C 118 (2014) 7591–7598

[68] Bing Gao, Qiqi Yan, Yu Tong, Xianghua Zhang, Hongli Ma, Jean-luc Adam, Jing Ren and Guorong Chen, J. Lumin. 143 (2013) 181–184

[69] Woan-Jen Yang, Liyang Luo, Teng-Ming Chen and Niann-Shia Wang, Chem. Mater. 17 (2005) 3883–3888

Accepted manuscript

ACCEPTED MANUSCRIPT

Figure Captions

Fig. 1: XRD pattern of the our data compared with $Ba_2Y(BO_3)_2Cl$

Fig. 2: Coordination of cations in $Ba_2Y(BO_3)_2Cl$

Fig. 3: SEM image of $Ba_2Y(BO_3)_2Cl$ phosphor

Fig. 4: PLE (a, λ_{em} = 420 nm) and PL (b-e, λ_{ex} = 360 nm) spectra of $Ba_{2(1-x)}Y(BO_3)_2Cl:2xCe^{3+}$ (*x*= 1 - 4 mol%) phosphors. Inset shows variation of emission intensity with activator concentration.

Fig. 5: PLE (a, λ_{em} = 420 nm) and PL (b, λ_{ex} = 360 nm) spectra of $Ba_{2(1-x)}Y(BO_3)_2Cl:2xCe^{3+}$ (x = 2 mol%) phosphors.

Fig. 6: PLE (a, λem = 420 nm and b, 1060 nm) and PL (c-i, λex = 362 nm) spectra of the $Ba_2Y(BO_3)_2Cl:Ce^{3+}$(2 mol%), Nd^{3+}(1-12 mol%) samples.

Fig. 7: Typical PLE (a, λem = 1060 nm) and PL (b, λem = 823 nm) spectra of Nd^{3+} in $Ba_2Y(BO_3)_2Cl$ host

Fig. 8: The dependence of Ce^{3+} (420 nm) and Nd^{3+} (1060 nm) emission intensity on the doping concentration of Nd^{3+} in $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$.

Ce^{3+}(2 mol%), Nd^{3+}(1-12 mol%)

Fig. 9: Decay curves corresponding to Ce^{3+}: 5d-4f transition in $Ba_2Y(BO_3)_2Cl:Ce^{3+},Nd^{3+}$ with different Nd^{3+} concentrations under 374 nm excitation.

Fig. 10: Schematic energy level diagram of Ce^{3+} and Nd^{3+} in the $Ba_2Y(BO_3)_2Cl$ phosphor, showing the energy transfer mechanism for the NIR luminescence under 362 nm excitation.

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 1. Average lifetime (τ_{avg}) (ns) and energy transfer efficiency (η_{ETE}) (%) of Ce^{3+} emission in singly doped, and Ce^{3+}, Nd^{3+} co-doped $Ba_2Y(BO_3)_2Cl$ phosphor.

x mol% Nd^{3+}	a_1	τ_1 (ns)	a_2	τ_2 (ns)	τ_{avg} (ns)	η_{ETE} (%)
0	0.0311	7.44	0.9688	25.5	25.37	0
1	0.0090	2.64	0.9909	11.8	11.79	53.5
3	0.0994	2.32	0.9005	9.72	9.53	62.4
6	0.1027	1.34	0.8972	6.22	6.10	75.9
9	0.0161	1.05	0.9838	3.86	3.84	84.8

Accepted manuscript

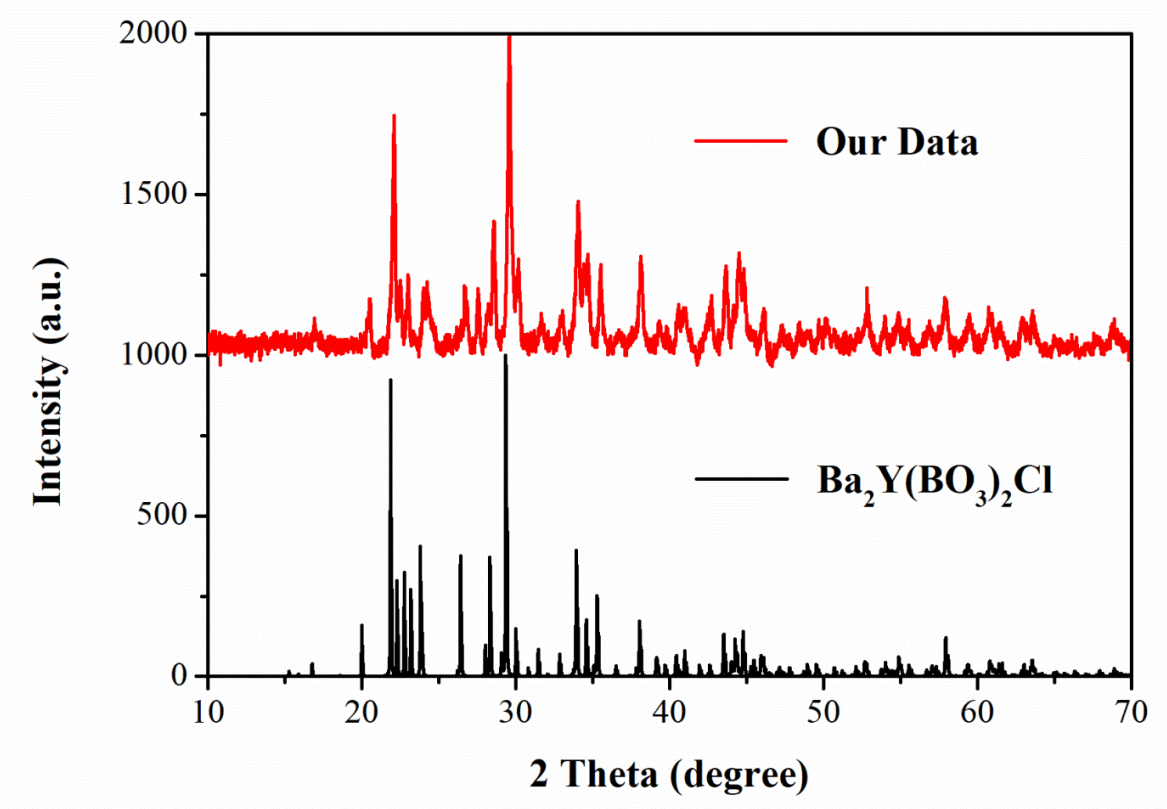


Fig. 1

Accepted manuscript

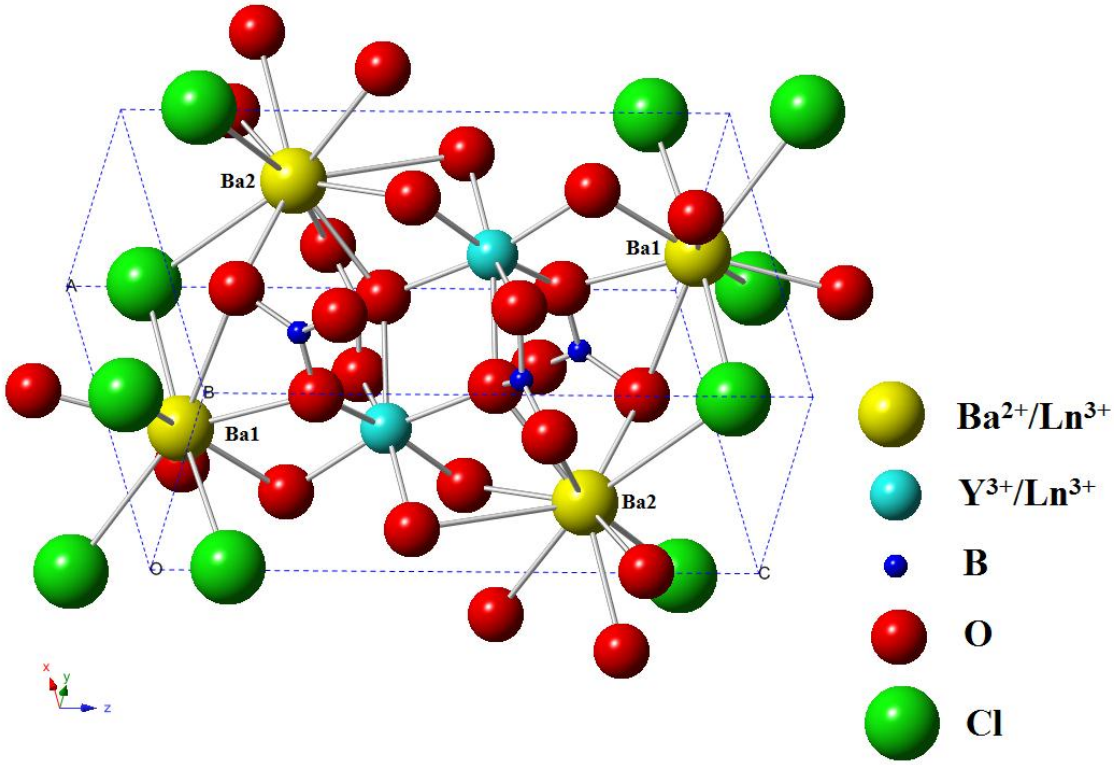


Fig. 2

Accepted ma

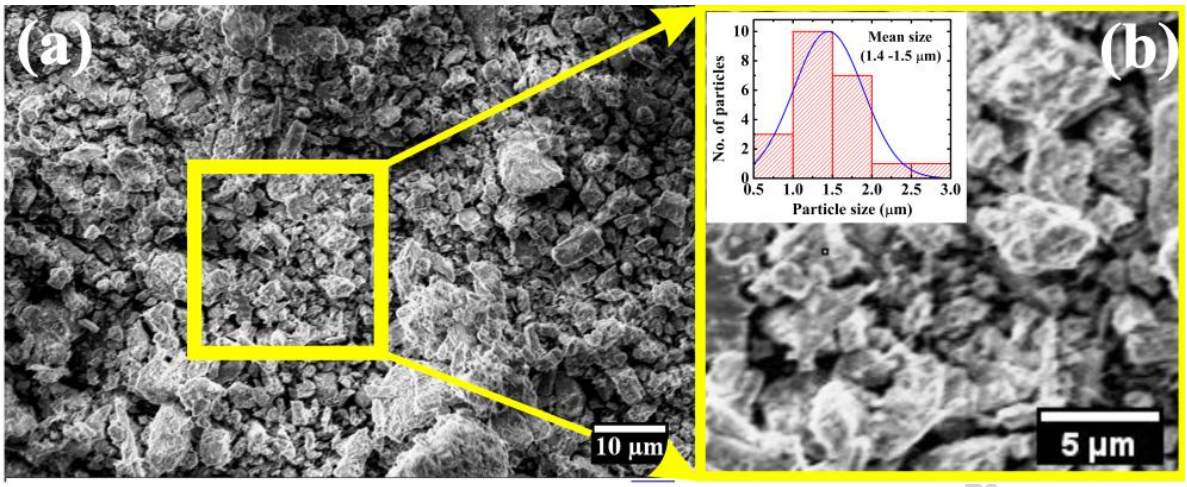


Fig. 3

Accepted manuscript

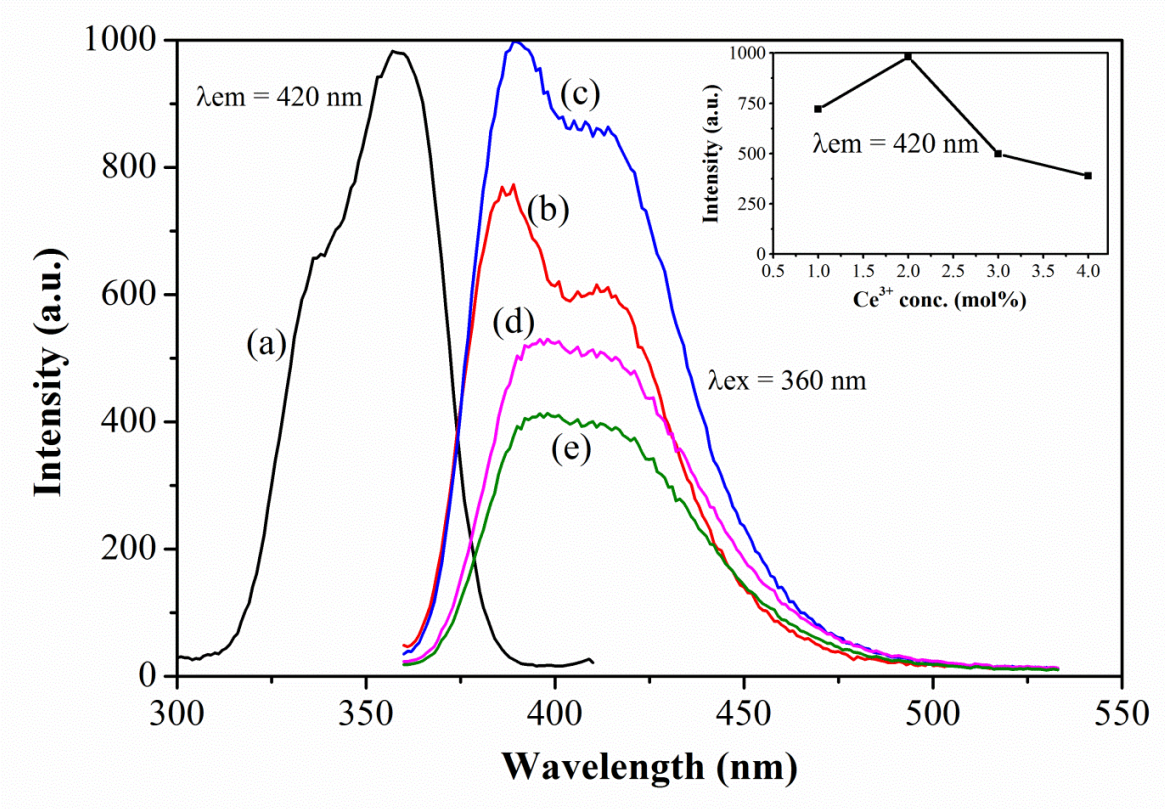


Fig. 4

Accepted manuscript

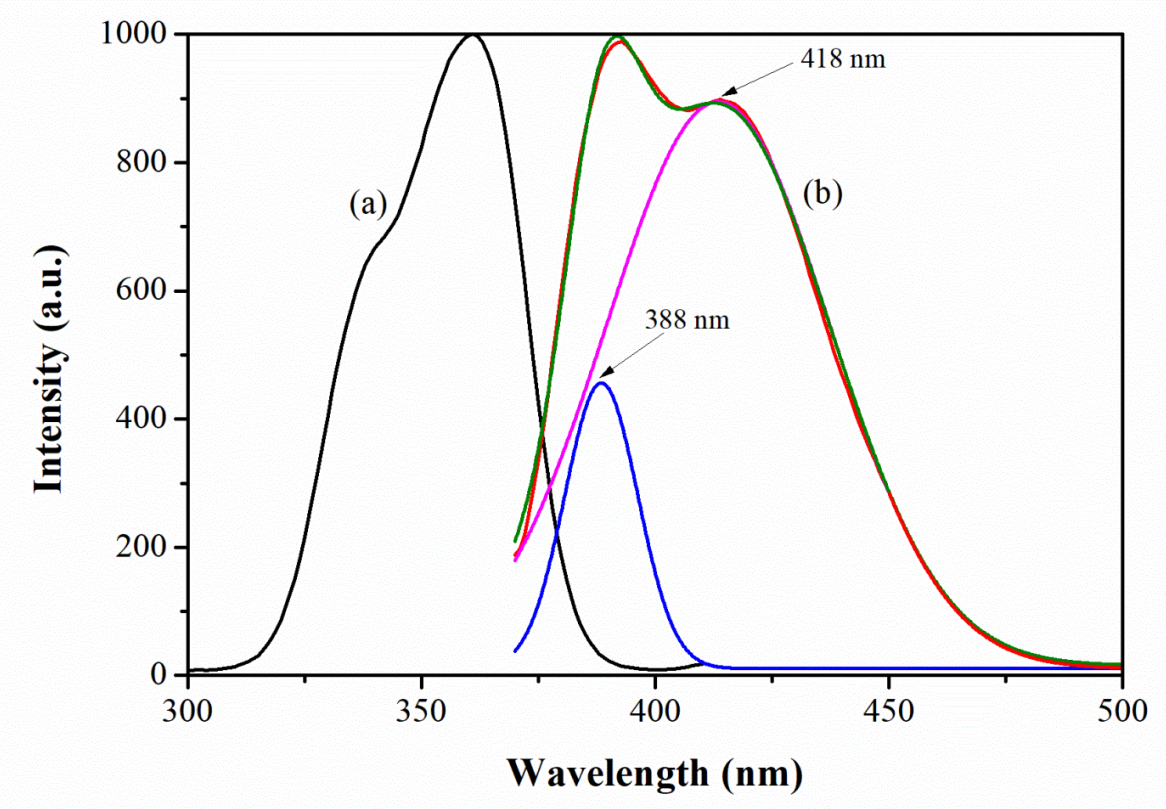


Fig. 5

Accepted ma

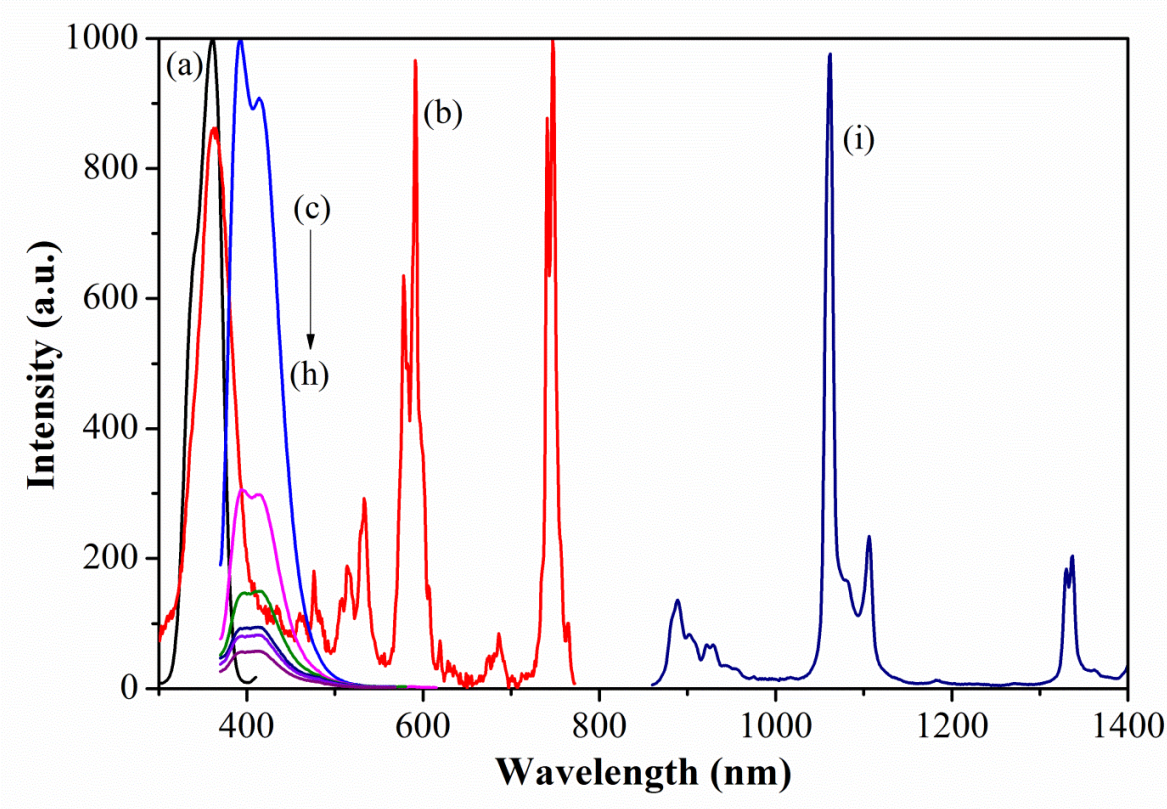


Fig. 6

Accepted manuscript

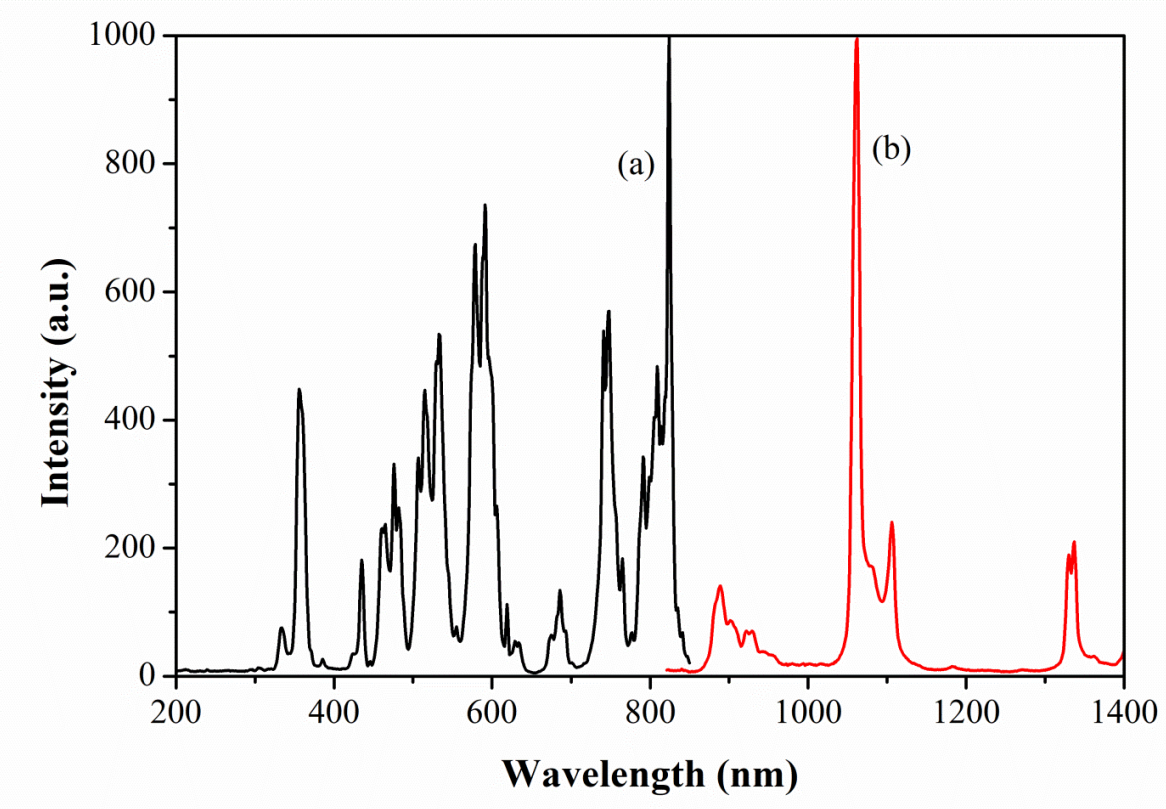


Fig. 7

Accepted manuscript

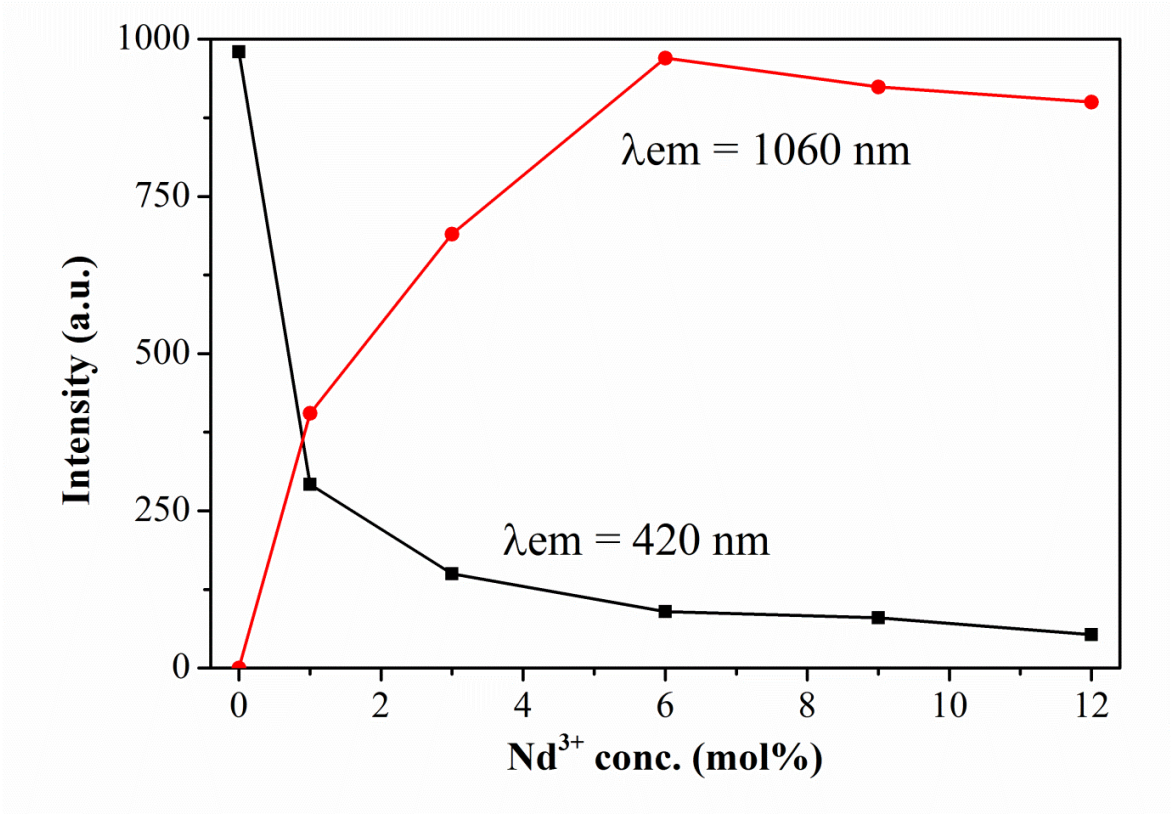


Fig. 8

Accepted ma

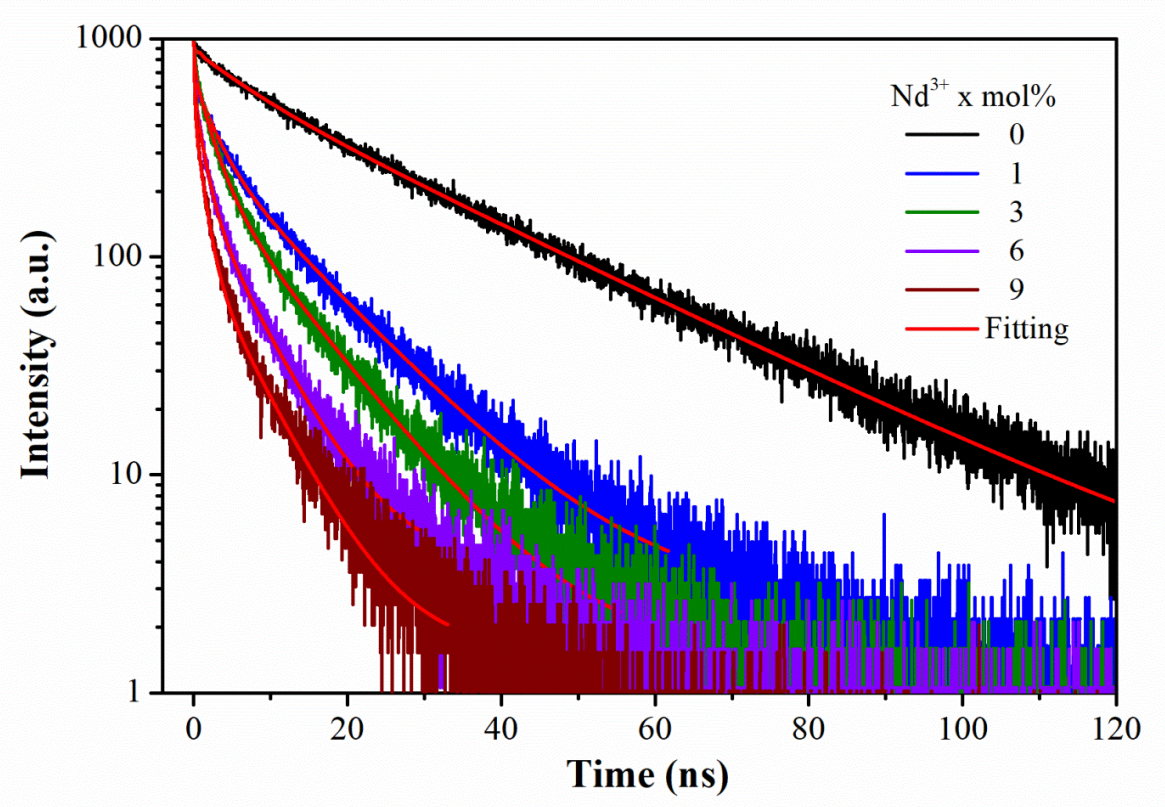


Fig. 9

Accepted manuscript

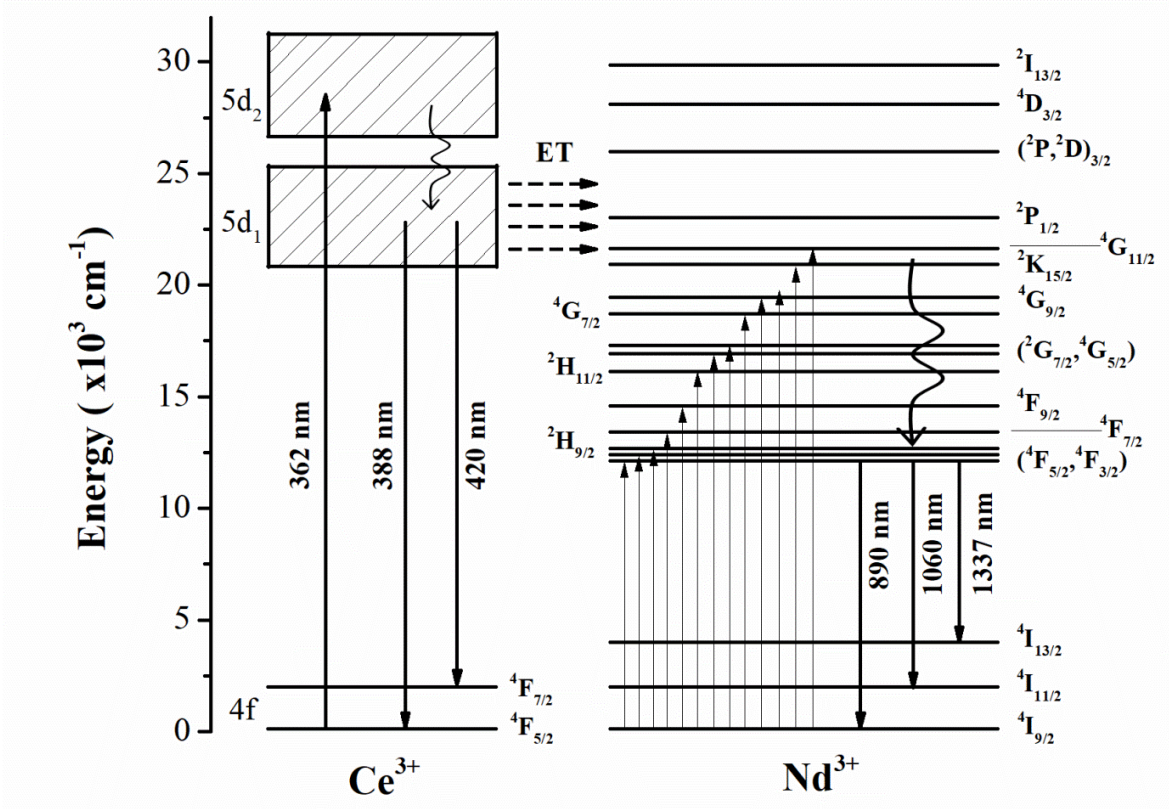


Fig. 10

Accepted manuscript

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 05 | May-2018 www.irjet.net p-ISSN: 2395-0072

Simulation of Inductively Coupled Plasma of $Ar/C_2H_2/CH_4/H_2$ gas mixture in PECVD reactor and calculating the reactor efficiency

Manish Kumar[1], Sagar Khanna[2], Suresh C Sharma[3]

[1,2]M.Tech Student, Dept. of Applied Physics, Delhi Technological University, Delhi, India,
[3]Professor and HOD, Dept. of Applied Physics, Delhi Technological University, Delhi, India.

Abstract - *Inductively Coupled Plasma based CVD technique plays an important role in thin film deposition and formation of carbon nanostructures at a temperature below 600°C (1). ICP involves generation of plasma by applying radio frequency-based power (2). During the formation of plasma, various losses occur in the system which is difficult to measure in experimental conditions precisely. Simulation model helps in mimicking actual experimental condition for calculating these losses and ultimately the overall efficiency of the PECVD system. COMSOL Multiphysics can be used as a tool to measure the losses and by careful manipulation of various parameters, the optimized set of parameters can be adjusted for more economical and safer plasma generation for thin films and carbon Nanostructure synthesis.*

Key Words: Inductively Coupled Plasma, COMSOL Multiphysics, Thin film deposition, carbon Nanostructures

1. INTRODUCTION

Carbon Vapor Deposition methods have been the major source for the thin film depositions at Nanoscale and formation of carbon structure at Industrial level because of their ability to produce these structures at 500 to 1100° C, which makes it suitable for working with delicate substrates. Typically, Radiofrequency plasma conditions are suitable for breaking gases into radicals and ions for deposition purpose. A commercial PECVD reactor comprises of a spiral antenna having number of turns on top of the chamber and consists of a generator operating at Radiofrequency of 13.56 MHz providing power ranging from 1200 to 1500 W (3). Practically it is inconvenient to study the PECVD system at every time and measure various working parameters of it, so situational models are important in analyzing the system at each step. COMSOL Multiphysics 5.2a is an elemental analysis tool or solver which provides conventional physics-based systems available at one platform using solution of partial differential equations. It can be incorporated to design the same practical PECVD system and analyze various parameters with simple GUI. As practically it is impossible and dangerous to measure losses in a PECVD unit, here COMSOL is handy in calculating various system parameters along with various losses in the system. By calculating various losses in a PECVD system, the overall efficiency of PECVD can be easily calculated using COMSOL, which is practically not possible in the experimental system.

2. MODEL

COMSOL Multiphysics 5.2a is used to model a two-dimensional PECVD model of plasma of complex $Ar/CH_4/C_2H_2/H_2$ gas mixtures at low-pressure of 20mTorr and low temperature of 300K to simulate a practical PECVD system and calculate the various losses in the system and ultimately calculate the overall efficiency of PECVD system.

2.1 Geometry

For analysis purpose of a 3D plasma chamber, a 2D axis-symmetric geometry is considered.

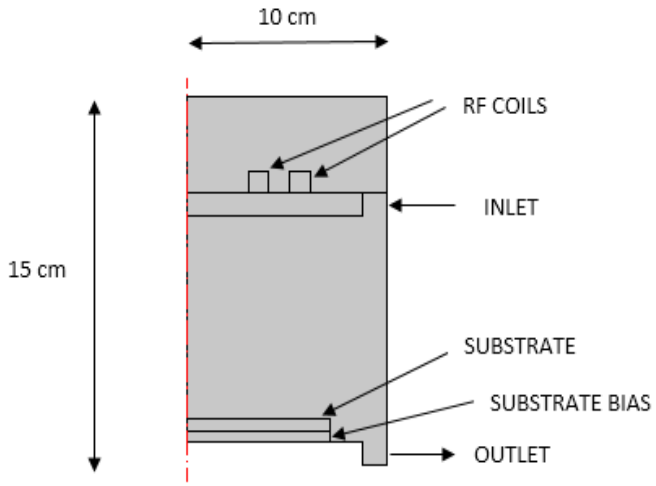


Fig- 1. 2D Axis Symmetric Model of 3D PECVD System

The plasma reactor chamber consists of a metal cylinder of radius 10 cm and height 15 cm as shown in Fig-1. A substrate platform of radius 7 cm is placed 8 cm below the chamber top. A four-turn spiral antenna (inductive coil) named Coil 1 is located at the top of the chamber and is connected to a 13.56 MHz generator providing power of1200 W. This RF coil is the main source for ICP plasma source. The antenna configuration with only four coil turns was chosen intentionally to generate low-temperature plasmas with non-uniform density along the radial direction.

The model incorporates Inductively coupled Plasma module under plasma physics of COMSOL. This interface couples the Drift-Diffusion, Heavy Species Transport, and magnetic field interfaces into an integrated Multiphysics interface.

2.2 Reactions and Materials

Table -1: Reactions considered for the Simulation Model

Electron Impact reactions				
S No	Reaction	Type	Threshold(eV)	Ref.
1	e+Ar=>e+Ar	Elastic	0	4
2	e+Ar=>e+Ars	Excitation	11.5	4
3	e+Ars=>e+Ar	Super elastic	-11.5	4
4	e+Ar=>2e+Ar+	Ionization	15.8	4
5	e+Ars=>2e+Ar+	Ionization	4.24	4
6	Ars+Ars=>e+Ar+Ar+	Penning Ionization	-	4
7	Ars+Ar=>Ar+Ar	Metastable quenching	-	4

Plasma chamber Wall Surface reaction			
No	Reaction	Sticking Coefficient	Ref.
8	Ars=>Ar	1	5
9	Ars+=>Ar	1	5

Reactions			
No	Reaction	Rate Coefficients (m^3/(s.mol))	Ref.
10	Ar++C2H2=>C2H2++Ar	25.29 x10^7	6,7
11	Ar++H2=>Ar+H2+	0.448 x10^8	6,7
12	C2H2++H2=>C2H3++H	0.1656 x10^6	6,7
13	H2++H=>H++H2	1.062 x10^7	6,7
14	C2H2+H+=>C2H++H2	0.713 x10^8	6,7
15	C2H3++H=>C2H2++H2	1.126 x10^6	6,7
16	C2H++H2=>C2H2++H	0.182 x10^8	6,7
17	C2H2++H+=>C2H2++H2	0.713x10^8	6,7
18	H++CH4=>CH3++H2	0.385x10^8	8
19	H2++CH4=>CH3++H2+H	0.379x10^8	8
20	CH3++H=>CH2++H2	1.55x10^7	8
21	CH2++H=>CH++H2	0.165x10^8	8
22	CH++H=>C++H2	1.242x10^7	8

Table -1: Materials used for the simulation model

Medium	Electrical conductivity	Relative Permittivity	Relative Permeability
RF coils (Coil 1)	6*10^7 S/m	1	1
Dielectrics /Insulator	0	1	4.2

2.3 Losses in PECVD system



Fig- 2 Flow of losses in PECVD

When a time-varying electric current is applied to the RF coils at 13.56 MHz and 1200 W power source, it produces a time-varying magnetic field in the plasma chamber filled with gases. As per Lenz's law, a time-varying electric current flows inside the chamber gases in opposite direction to oppose the magnetic field. This electric current in the gas chamber ionizes the Argon gas to form the plasma by transformer action.

The losses due to transformer action involve ohmic losses in the coils of PECVD and Iron losses which include Hysteresis and eddy current loss (9). The eddy current loss can be neglected as the coils are laminated cores. The constant Hysteresis loss is calculated as

$W_h = K_h B max^{1.6} f V$ (watts)

where, K_h = Steinmetz hysteresis constant, V = volume of the core in m^3, f=13.56 MHz, Total volume of Coil= 30.4 x 10^{-6} m^3, Maximum Flux Density= 0.01 T and K_h = 100

Total Hysteresis loss= 2 x W_h =26.82 W

Due to ionization of gases, the formation of plasma takes place breaking constituent gases in various Ions, radicals, and neutrals. It involves electron-impact reaction, various inter-species reactions, and reactions with plasma wall as specified in Table 1.

During plasma formation mainly two types of losses take place namely Capacitive power deposition loss and Collisional Power Loss. Another loss is volumetric losses due to the electric and magnetic field which is of the order of 10^{-14} and can be neglected.

The overall efficiency of a PECVD unit can be calculated using the formula

$$Efficiency, \eta = 1 - \frac{Total\ losses\ (Energy)}{Input\ (Energy)}$$

3. RESULTS

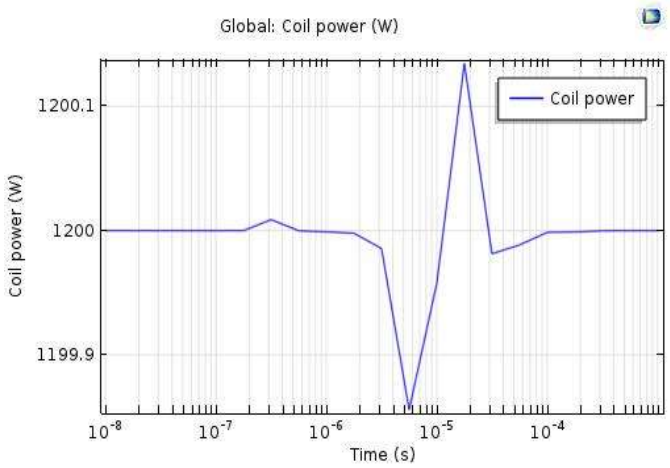


Fig- 3 Coil Power

Initially, the power dissipated is all dissipated in the coil (~1200 W). After about 1 microsecond, the plasma ignition begins and as the neutral gas atoms split into electrons and ions, the electrons begin to absorb more and more power. At this time a countercurrent start flowing in the coil as per Lenz's law. Over a period of 2 microseconds, the system restores again by increasing its power intake up to 1300 W for few microseconds. The plasma goes from absorbing no power to absorbing around 1300 W and restored back to 1200W as shown in Fig- 3

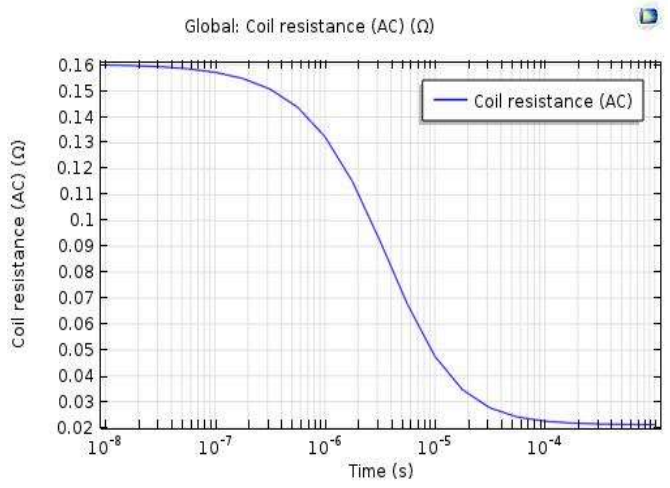


Fig- 4 Coil Resistance

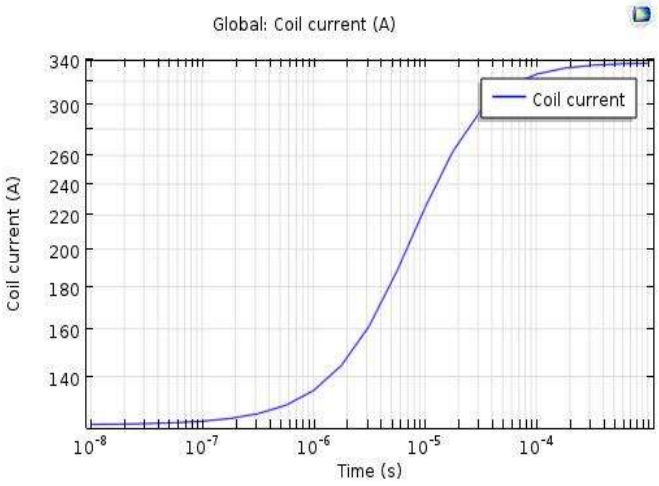


Fig- 5 Coil Current

The resistance of the coil decreases by a little less than a factor of 8 from 0.16 to 0.02 when the plasma is on as shown in Fig- 4. There is a substantial opposing current induced back into the coil from the plasma. The coil current increases from 135 A to 350 A after few microseconds when plasma is generated as shown in Fig- 5. The electric potential applied across the coil needs to increase in order to maintain the same total current

Using the coil current and coil resistance plot data the ohmic ($I^2 R$) loss plot is shown in Fig- 6

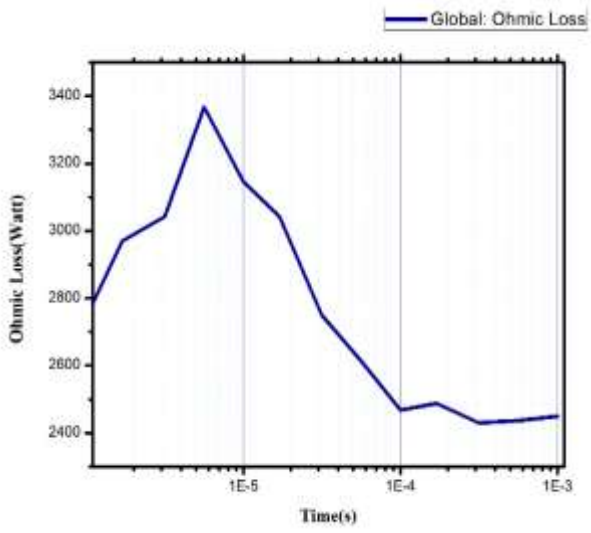


Fig- 6 Ohmic Loss

Using Trapezoidal rule in the curve of Fig.6 the ohmic loss is calculated to be 57.7 W.

Fig- 7 shows the energy loss due to capacitive power deposition and it gradually reduces to zero just after plasma starts to build up at microseconds and is calculated to be 2.2W.

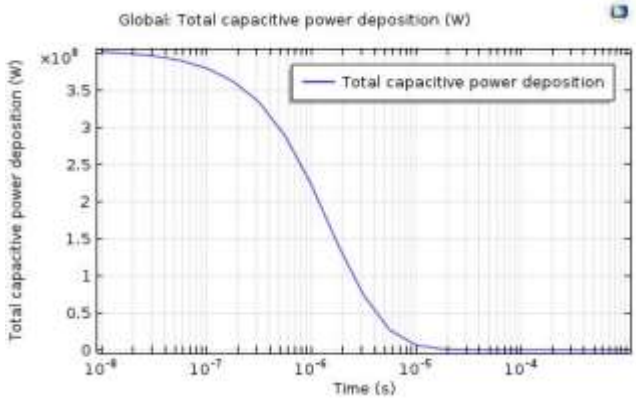


Fig- 7 Capacitive Power Deposition

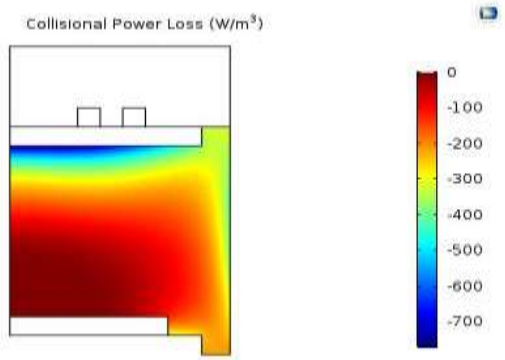


Fig- 8 Collision Power Loss

The per unit volume Collisional Power Loss is shown in Fig- 8 and is maximum near the coil and Plasma chamber interface with a value of 733 W/m^3. So, for total volume, it is calculated to be 2.47 W.

The initial electron density was considered to be 10^{12} /m^3. After plasma generation finally, the maximum electron density reaches up to a value of 6.06x10^{12} /m3 around the walls of plasma chamber due to various electron impact reactions as shown in Fig-9.

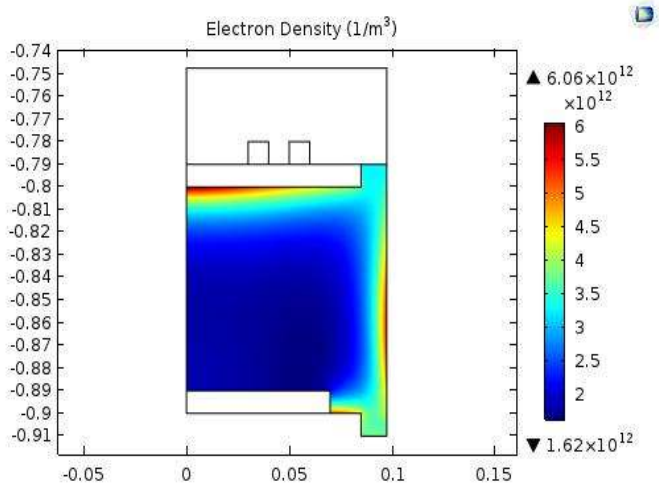


Fig- 9 Electron Density

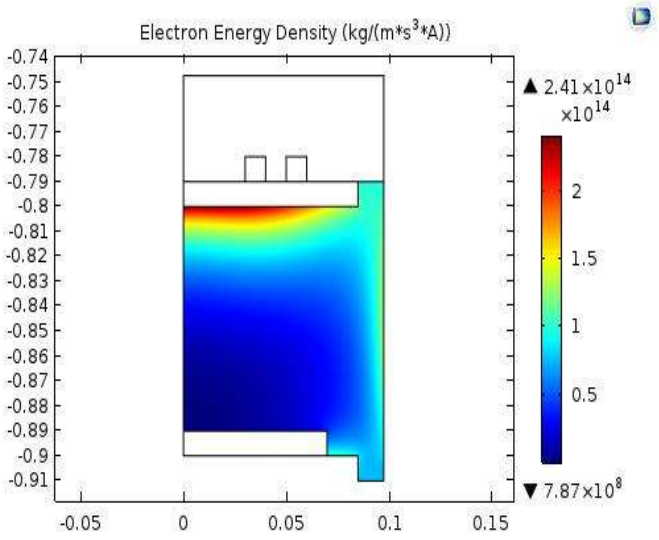


Fig- 10 Electron Energy Density

Fig- 10 shows the electron energy density inside the chamber reached up to a value of 2.41x10^{14}.

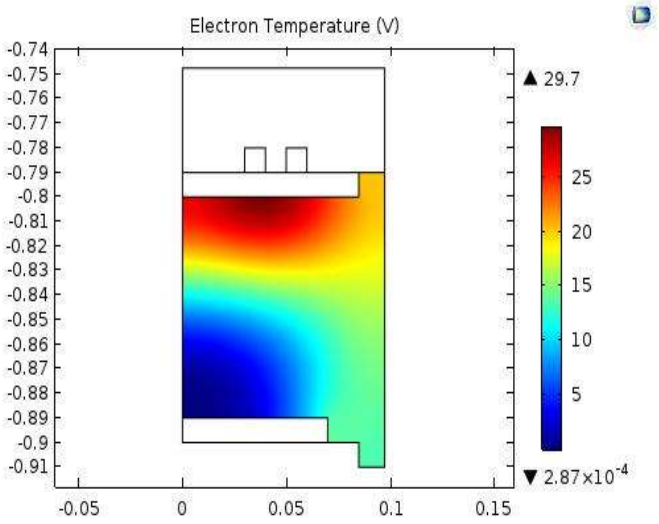


Fig- 11 Electron Temperature

The plasma generation in the chamber causes an increase in the electron temperature up to a maximum value of 29.7(in the unit of volts) which is equivalent to 11000 K which is shown in Fig- 11.

The electric potential generated inside the chamber due to plasma action is found to be 4.81x10^5 and is maximum in the center of the chamber.

Total loss in the system is calculated to be 89.19 W giving an overall efficiency of 92.51 %. Which mainly comprises ohmic loss in the coil and collisional power loss in the plasma chamber.

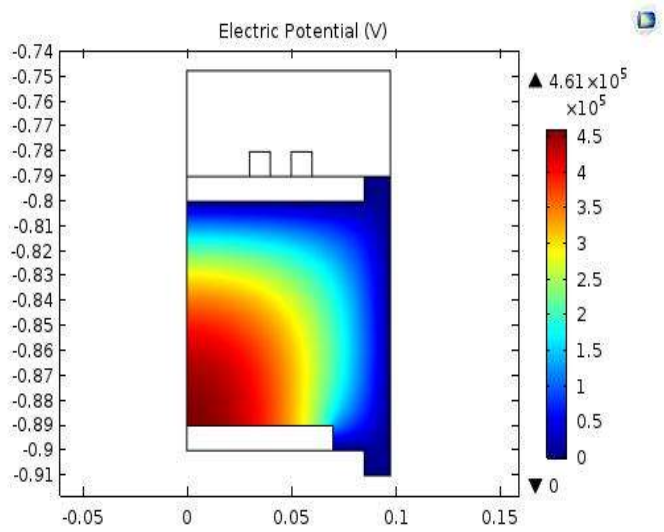


Fig- 12 Electric Potential

4. CONCLUSION

This paper presents the simulation results of low-pressure Inductively coupled RF plasmas in Ar/CH4/C2H2/H2. The calculation is performed by COMSOL Multiphysics. It is found that total losses in our PECVD unit are 89.19 W and the efficiency is calculated to be 92.51 % and mainly depends upon the type of material taken for coils, the operating parameters of coils. Also, the losses due to plasma formation depend upon the operating conditions of plasma chamber as initial electron density, temperature, pressure of chamber types of constituent gases.

REFERENCES

[1] B. Purna Chandra Rao, R. Maheswaran, Shivaraman Ramaswamy, Ojas, "Low-Temperature Growth of Carbon Nanostructures by Radio Frequency-Plasma Enhanced Chemical Vapor Deposition, 2009.

[2] Kostya (Ken) Ostrikov,* Hyun-Jin Yoona, Amanda E. Rider, Sergey V. Vladimirov, "Two-Dimensional Simulation of Nanoassembly Precursor Species in Ar/H2/C2H2 Reactive Plasmas, 2007.

[3] M. Keidar, A. M. Waas, Nanotechnology 2004, 15, 1571.

[4] 1. G.J.M. Hagelaar and L.C. Pitchford, "Solving the Boltzmann Equation to Obtain Electron Transport Coefficients and Rate Coefficients for Fluid Models," Plasma Sources Sci. Technol, vol. 14, pp. 722–733, 2005.

[5] D.P. Lymberopolous and D.J. Economou, "Two-Dimensional Self-Consistent Radio Frequency Plasma Simulations Relevant to the Gaseous Electronics Conference RF Reference Cell," J. Res. Natl. Inst. Stand. Technol., vol. 100, p. 473, 1995.

[6] Denysenko I B, Xu S, Long J D, Rutkevych P P, Azarenkov N A and Ostrikov K," J. Appl. Phys. 95 2713–24", 2004.

[7] Herrebout D, Bogaerts A, Yan M, Gijbels R, Goedheer W and Dekempeneer E," J. Appl. Phys. 90 570", 2008.

[8] Morrison N A, William C, and Milne W," J. Appl. Phys. 94 7031–43", 2001.

[9] Losses in Transformer Proceedings of the American Institute of Electrical Engineers.

J. Inst. Eng. India Ser. C
https://doi.org/10.1007/s40032-018-0454-6

REVIEW PAPER

Solar Energy a Path to India's Prosperity

Yogender Pal Chandra[1] · Arashdeep Singh[2] · Vikas Kannojiya[3] · J. P. Kesari[4]

Received: 9 March 2017 / Accepted: 29 March 2018
© The Institution of Engineers (India) 2018

Abstract Solar energy technology has grabbed a worldwide interest and attention these days. India also, having a huge solar influx and potential, is not falling back to feed its energy demand through non-conventional energy sources such as concentrating solar power (CSP) and photovoltaic (PV). This work will try to add some comprehensive insight on solar energy framework, policy, outlook and socio-economic challenges of India. This includes its prominent areas of working such as grid independent and 'utility-scale' power production using CSP or PV power plants, rural as well as urban electrification using PV, solar powered public transportation systems, solar power in agrarian society—water pumping, irrigation, waste management and so on and so forth. Despite the fact that, a vast legion of furtherance and advancement has been done during the last decade of solar energy maturation and proliferation, improvements could be suggested so as to augment the solar energy usage in contrast to conventional energy sources in India.

Keywords Renewable energy · Solar energy · JNNSM

✉ Yogender Pal Chandra
yogenderpal.chandra@fs.cvut.cz;
yogender027mae@gmail.com

[1] University Centre for Energy Efficient Buildings (UCEEB), Czech Technical University in Prague (CVUT), Prague, Czech Republic

[2] Mechanical Engineering Department, Thapar University, Patiala, India

[3] Department of Mechanical Engineering, Indian Institute of Technology Roorkee, Roorkee, India

[4] Department of Mechanical Engineering, Delhi Technological University (DTU), Delhi, India

Introduction

India has tremendous energy potential and despite finding difficulties, it is coping with the increase in demand to meet those needs through conventional as well as renewable energy sources of power generation. Electricity demand in India has been increasing at one of the fastest rates in the world due to population growth and economic development. As a matter of fact, according to census of India, only 55.5% of rural households had access to electricity in 2011. In another words, only 93.12 million households out of 167.8 million rural households were electrified in 2011. Accordingly, as per statistics, in few states the rural household electrification during 2011 was alarmingly low, namely, Bihar (10.4%), Assam (28.4%), Uttar Pradesh (23.8%), Odisha (35.6), Jharkhand (32.3%), and West Bengal (40.3%) [1, 2]. Hence there was a huge lacuna between demand and supply of electricity for rural population during 2011. Contrasting this, Ministry of Power, GOI has taken some profound initiative and had sanctioned around 921 projects to electrify 1,21,225 un-electrified villages, together with intensive electrification of 5,92,979 partially electrified villages and had provided free electricity connections to 397.45 lakh BPL rural households [2]. With regards to this, electrification of rural areas has been transcended so far at an exceedingly swift rate. This includes out-breaking statistics for Bihar (95%), Assam (96.8%), Uttar Pradesh (98.7%), Odisha (91.9%), Jharkhand (92.8%), and West Bengal (99.99%) [2]. To illustrate the above mentioned reforms in rural areas were resulted due to the announcement and propagation of Deendayal Upadhyay Gram Jyoti Yojana (DDUGJY) by GOI with chief modifications in preexisting Rajiv Gandhi Grameen Vidyutikaran Yojna (RGGVY). The aims of the policy were strictly complied and include solidification and

Published online: 10 May 2018

intensification of sub-transmission and distribution infrastructure in rural areas including metering of distribution transformers / feeders / consumers, rural electrification as per CCEA approval dated 1 August 2013 for completion of the targets laid down under RGGVY for 12th and 13th Plan [3]. Furthermore, though India's growth in energy is apparently upright, however to uphold this growth in the recent years has put a significant pressure on conventional energy sources or fossil fuels. Henceforth, there immerges a considerably great need of affordable and clean source of energy with a view to satisfy overgrown and insatiable energy demands. Correspondingly, with this intensified energy demands, which actually happened within a decade or so, achieving renewable and sustainable supply for the future has become a great challenge. As a matter of fact, India is endowed with vast solar energy potential. About 5000 trillion kWh per year energy is incident over India's land area, in another words, solar energy ranging around 4 to 7 kWh/m^2 per day is incident on India's soil with global radiation ranges from 1700 to 2300 kWh/m^2 annually, Fig. 1 explains the same [4, 5]. Correspondingly, due to maturation and uprising in solar energy technology, solar power generation has become inexpensive over the time. Solar energy has turned into an economical power source in long-run due to its profound benefits in contrast to conventional energy sources. All the same, only fraction of the India's total power consumption comes from solar energy. For instance, out of total installed capacity of 303,083.21 MW during May 2016, renewable energy source (RES) amounts an insufficient value of 42,849.38 MW out of which solar energy contribution was merely 6762.85 MW. In view of this, at RE-INVEST 2015 which was the first global investors meet in the area of renewable energy, conducted at New Delhi, 293 companies committed to generate 266 GW of renewable energy within period of next five years by building and setting up power plants based on RES and also building infrastructure to manufacture goods and services related to RES, thus facilitating youth employment and awareness in the same [6]. Renewable energy also has the front hand thus empowering people at grass root level. The technology for setting up solar energy systems is widely available in India, it is required to bring them to serve the nation's need.

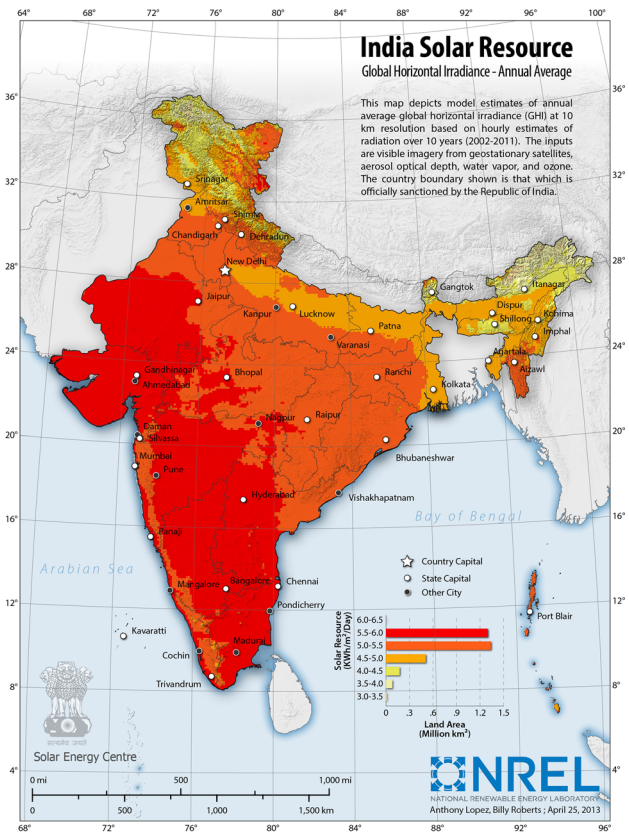


Fig. 1 Global horizontal irradiance (GHI) of India [5]

India's Global Initiative

A major contribution by India to address its energy security with the help of RES, more particularly solar energy, as well as the global climate change issue was well marked in January 2010 by the former Prime Minister Dr. Manmohan Singh under the name of Jawaharlal Nehru National Solar Mission (JNNSM) also known as National Solar Mission. The aim of the project is to generate 20 GW of power by 2022, which was later on augmented to 100 GW. Table 1 illustrates the road map of JNNSM work plan for the next upcoming year. Furthermore, in 2015 union budget of India, some serious and ambitious goals were set that would be followed very strictly and aimed at [7]:

- Reducing cost of power production by solar energy to the ground level.
- Energetic R&D.
- Domestic production of raw materials through rigorous implementation of long term goals with a large scale deployment of goals.

So far, 99 projects with a cumulative capacity of 252 MW have been installed under JNNSM. The state-wise break-up of the number of projects and capacity installed as given in Table 2 [7].

Solar Schemes

With the constitution of JNNSM a rigorous task objectives were defined to use solar energy under various schemes as explained here.

Table 1 Proposed road map of JNNSM [7]

Application segment	Target phase—1 2010–2013	Target phase—2 2013–2017	Target phase 3 2017–2022
Grid solar power, MW	1100	4000–10,000	20,000
Off grid solar power, MW	200	1000	2000
Solar thermal, million sq. meter	7	15	20

Table 2 Status of JNNSM [7]

State	NVVN, Batch 1		RPSSGP, IRDEA		Total	
	No.	MW	No.	MW	No.	MW
A.P.	2	10	10	9.75	12	19.75
Chhattisgarh			2	4	2	4
Haryana			8	7.8	8	7.8
Jharkhand			6	12	6	12
Maharashtra			3	5	6	16
Orissa	1	5	7	7	8	12
Punjab			5	6	6	8
Rajasthan	20	100	10	10	38	147.5
T.N.	1	5	5	5	6	10
U.P	1	5	3	5	4	10
Uttrakhand			3	5	3	5
Total	25	125	62	76.55	99	252

Solar Park Scheme

Developing and setting up of 25 solar parks around the country, each with a capacity of around 500 MW and a broad objective of reaching 20,000 MW of power capacity within next five years. The extensive vigour of this scheme is to encourage and facilitate project developers and investors, reduce cost in solar equipment procurement with impetus in technical improvement and R&D [7].

CPSU's Scheme

The Central Public Sector Undertaking (CPSUs) has initiated for implementation and setting up of 1000 MW grid connected Solar PV power plants within a span of three years, that is, 2014–2015 to 2016–2017 [7].

Defence Scheme (Rooftop Projects)

This scheme is aimed to promote rooftop area of defence sector/paramilitary force utilization to generate electricity using PV panels. Both grids—connected as well as well as off—grid projects are taken into consideration. Objectives of this scheme is targeting for more than 300 MW of power production, within the establishment of Ministry of Defence, namely, area of Army, Navy, Defence Laboratories, Air Force, Para Military Forces and Ordinance Factories. This scheme is also believed to boost indigenously made solar cells and modules [7].

CSP Technology: India's Perspective

There comes two vogue classes of technologies used to conserve the solar energy namely concentrated solar power (CSP) and photovoltaic (PV). While the latter employs the use of solar photovoltaic cells to produce electricity through photoelectric effect. The former employs various methods of heat capturing devices and structures, thus finally using that thermal energy into some power producing thermal cycles, more preferably rankine cycle.

Concentrating solar power has been a subject of great interest worldwide for several decades with Spain in its leading rein [9]. Concentrating solar energy employs a complicated scheme of mirrors with a sun tracking arrangement, heat transfer fluids to transport the necessary thermal energy to the power block where it can be fed to power turbine to generate electricity through some power cycle, rankine cycle more preferably.

CSP Plants Sanctioned in JNNSM, Phase I, Batch-1

CSP projects sanctioned and their location in Phase-I, Batch-1 are shown in Table 2.

While, the projects named Rajasthan Sun Technique under Reliance Power, Godavari Green and MEIL Green Power are already commissioned in 2014, 13, 14 respectively, rest of the projects are in their nascent stages [8].

Target against Achievement during JNNSM Phase-I: A Feasibility Report

First demonstration project held by ACME Solar in Rajasthan was 2.5 MW solar tower based power plant. It was 46 m high and was commissioned in 2011 after two years of its long construction. The major issue regarding this plant was, it was built on solid grounds of

NASA satellite proposed DNI availability at the site. However, due to substantial discrepancies in proposed availability of DNI and actual DNI (which was actually considerably low), impacted the financial feasibility of all round running of the plant. Due to low solar availability even in the perfect sunshine, CSP plant as a matter of fact was consuming even more energy than it was producing and henceforth it had to be shut down temporally until the glitches would be resolved. Figure 2 establishes the relation between target set in JNNSM, Phase I, Batch-1 and the output so far.

JNNSM Phase-II

CSP performance in JNNSM slumped terribly due to various techno-economic concerns, henceforth CSP share was reduced to 30% by MNRE in Phase-II and targets were distributed among centre and states. In other words, the proposed power production through CSP now has been reconsidered to be reduced to only 30% of the previous proposed assessment and that also is divided among state and centre in the ratio of 60 to 40%. This is due to the fact that there is still lack of experience in CSP to define any further course of action in Phase-II. However, Solar Energy Corporation of India (SECI) is expecting to propose two projects of CSP, each of 50 MW capacities, so CSP still not has been eradicated. On contrary, in Phase-II the proposed 1500 MW CSP project was dramatically changed to 750 MW PV based power plant. Table 3 illustrates the same [7].

In addition, as a state initiative, Gujarat has announced 968.5 MW solar power production out of which 690 MW has already been commissioned. Karnataka has contributed for 8 MW so far (Table 4). Rajasthan being in forefront has contributed 200 MW out which 100 MW is from ST and 100 MW from PV [7]. That is to say, JNNSM, Phase-II has already encouraged the states to take initiatives and as a result over 760 MW capacity projects are installed [7].

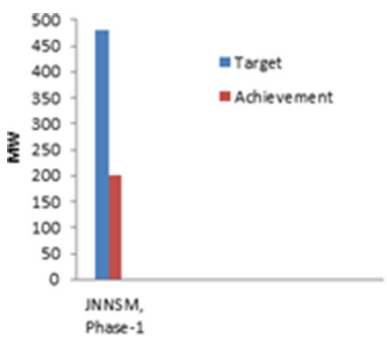


Fig. 2 Target achieved so far

Other Solar Thermal Technologies

Concentrated Solar Thermal Energy Technology based on Parabolic Dish Concentrators

The project is a cooperative effort between Megawatt Solutions Pvt. Ltd. and National Institute of Solar Energy (NISE) under MNRE's initiative to promote research, development and demonstration of indigenously developed renewable energy systems and technologies under a cost-sharing basis. It involves demonstration and evaluation of four interconnected dish concentrators each of 90 m^2 aperture area providing heated thermic fluids up to 400 °C [10, 11].

Solar Stirling Engine

The project has been taken up in collaboration with Oil and Natural Gas Commission (ONGC) Energy Centre (OEC). Three units of engines of 3 kW capacities each have been installed and commissioned in the campus. The objective of the project is to carryout long term performance evaluation under Indian conditions. The engines have been connected to the local grid and the electricity produced during the sunshine hours is being utilized in the Technical Block. The rated output of the facility is 9 kW (peak power) at solar insolation of 850 W/m^2 at 20 °C ambient temperature [10, 11].

Central Receiver Concentrated Solar Power Plant for Decentralized Power Generation

Sun Borne Energy with support from MNRE will be setting up a 1 MW CSP central tower pilot facility in partnership with NISE. The main objective is to develop the optimized designs of the heliostat field, volumetric air receiver and thermal storage, the three major components of a concentrated solar power (CSP) central receiver plant and also to develop the local sources for all the key components of the plant with a focus on lowering costs [10].

Solar Thermal Air-conditioning Systems

The project (100 kW cooling capacity) is being implemented by Thermax Limited at NISE with an objective to integrate parabolic trough collectors, vapour absorption machine (VAM) and appropriate thermal storage system for consistence performance of the system [7].

Table 3 CSP projects in India [8]

Project	Size and technology	Location
Lanco Solar 'Diwakar'	100 MW PTC	Askandra, Jaisalmer, Rajasthan
KVK Energy Ventures	100 MW PTC	Askandra, Jaisalmer, Rajasthan
Reliance Power	100 MW LFR	Dhusar, Jaisalmer, Rajasthan
Corporate Ispat Alloy/Abhijeet	50 MW PTC	Nokh, Jaisalmer, Rajasthan
Godawari Green	50 MW PTC	Nokh, Jaisalmer, Rajasthan
Aurum	20 MW PTC	Mitrala, Porbandar Gujarat
MEIL Green Power	50 MW PTC	Pamidi, Ananthapur, Andhra Pradesh

Table 4 State and centre involvement [7]

Technology	Total		Centre		State	
	Ratio, %	MW	Ratio, %	MW	Ratio, %	MW
Solar PV	70	6320	40	2520	60	3780
CSP	30	2700	40	1080	60	1620

Table 5 State involvement to PV in JNNSM, Phase-II [7]

State	Capacity, MW	Month of commissioning
A.P.	3	2011
Chhattisgarh	4	2011
Delhi	2	2010
Gujarat	51	2011
Haryana	2	2011
Karnataka	6	2009
Maharashtra	18	2011
Orissa	3	2011
Punjab	3	2009–2011
Rajasthan	43.5	2010–2011
T.N.	7	2010–2011
W.B.	1	2009
Total	143.5	

PV Technology: India's Perspective

Table 5 demonstrates state wise distribution of PV projects already commissioned by 2011. Grid connected PV projects may be either ground mounted or rooftop PV. Solar Energy Corporation of India (SECI) has been designated as the nodal agency for the implementation of 750 MW of solar PV projects under JNNSM Phase II, Batch-I, wherein SECI has been entrusted with the responsibility of projects selection, monitoring and timely execution, handling VGF funds and trading of the power generated. The 750 MW of projects have been selected through a transparent process of VGF-based reverse bidding, 50% of which have the mandate of domestic content requirement (DCR). The selected projects are likely to be commissioned by April, 2015. However, no status has been uploaded till now. Furthermore, MNRE plans setting up 25 solar parks, each with a capacity of 500 to 1000 MW; thereby targeting around 20,000 MW of solar power installed capacity. These solar parks will be put in place in a span of five years and the solar projects may then come up as per demand and interest shown by developers [9].

PV Rural Requirement in India

Cold Storage with Solar: Biomass Hybrid System

It is a project in partnership with TERI, Thermax Limited, NISE and CSIRO Australia with an objective to develop cold storage particularly in rural areas utilizing exhaust heat of biomass gasifier engine / solar scheffler dish [12].

Solar Powered Battery Well

The concept of solar powered battery well is similar to the village community water well. In this case, the solar energy is harvested and stored and inhabitants can charge devices or energy storage systems from energy stored in batteries for a fee. The installation and maintenance of this system can be done through a village cooperative arrangement so that it can be made accessible to inhabitants at a lower cost per watt [10].

Solar Powered Vehicles

For fast transportation of materials, crops and humans, rural areas now depend on vehicles that run on gas / patrol. These vehicles could be designed to run on solar power.

Solar Powered Food Preservation Facility

Farm products degrade in normal climatic conditions. Low temperatures and humidity allows these products to extend their usable life. Storages with compartmental design can be built in rural areas and their environment can be regulated with solar energy technology. These storages can be built through village cooperatives and then rented to farmers for their use. Food Preservation will greatly help to satisfy India's growing demands.

Solar Lighted Streets / Farm Lands

Solar lights are very commonly used in many metropolitan or other cities to provide safety and necessary illumination without any need to power connection from grid to them. Now there is an impending need for these solar lights in rural areas as well.

Solar Powered Rural Homes

The technologies that can be solarised at home include lights, phones, refrigerators, food cookers, heaters, television, air / water coolers, water well and fans etc.

Solar Powered Waste Management

Farms and farm homes produce waste which largely pollutes land, water, and air thereby becoming a health hazard, in addition to being an eye sore. Solar powered compressors or waste conversion to useful products would benefit immensely the rural communities.

People's Perception of Solar Energy

If solar energy provides such splendid benefits over conventional sources then why it contributes for such a small fraction in power sector of India. At a glance, solar energy as on April 11, 2017 is contributing only 3.12% to the state, central and private power sector in India, and this constitutes around 20% of the contribution provided by all the renewable energy sources. The reason for this slow solar power development directly doubts towards adverse impairments such as social, cultural, bad political regulation etc. Among these, social-techno aspects are the critical barriers as people have strong inclination towards status quo. However, this trend is gradually changing and people are becoming more aware towards non-conventional sources as a step towards conservation of fossils and the plant. In a survey conducted by WWF-India on the awareness of the current status of renewable energy in India, more than 90% of the respondents voted in the favour of solar energy that it has very good prospects, while comparatively lower, that is, 60% thought that waste to heat energy has high potential, refer Fig. 3. In addition, it is well known fact that customers don't demand technology they demand services. Henceforth, lightening application in the rural India has been promoted most using solar photo voltaic and the traditional Research Development Demonstration and Extension (RDDE) approach has been changed to market orientation and commercialisation. This has magnificently enhanced the outlook of suppliers and manufactures restricted not only to installation but also implementation of aftersales services for better customer services. Rural youths and locals are trained in the solar PV technology in the fields [12]:

- Roof top installation and commissioning
- Installation of inverters and controllers in rural primitive areas.
- After services such as repairmen and changing the defective parts.
- To be a self-dependent 'solar barefoot mechanic'.

Research Facility

National Solar Thermal Power Testing, Research and Simulation Facility envisage a grid connected Solar Thermal Power Plant of 1 MW capacity. This also includes a test set up that enables companies and research institutions to test the performance of different solar concentrators, coatings and materials, components and system for a solar thermal power plant. The project is being implemented by

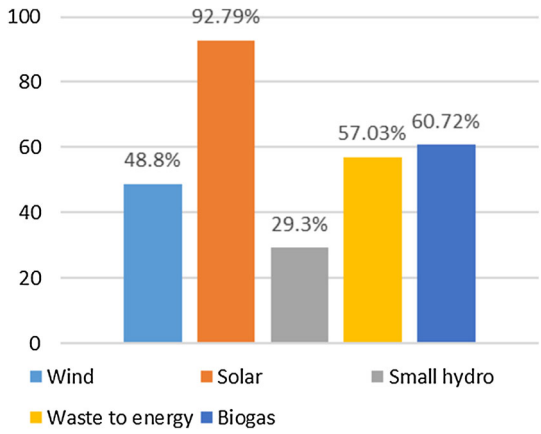


Fig. 3 Public acceptability of different renewable energy [12]

IIT Bombay and a consortium partners consisting of Tata Power, Tata Consulting Engineers, Larsen & Toubro, Clique, KIE Solatherm. This research facility is located at the campus of NISE, Haryana.

Investment Opportunity and Financing in Solar Projects and Research Facilities

Flamboyant stage persona and dramatic growth of nationwide solar policy called Jawaharlal Nehru National Solar Mission has facilitated the unilateral and bilateral ties of MNRE with number of Financial Institutions (FIs) such as World Bank, US Exim, JICA, EIB, ADB, AFD etc to avail line of credit. In addition to these FIs, to execute the plan of 175 GW of renewable energy, out of which 100 GW is dedicated to solar power, an investment of about $ 140 (or Rs. 8,50,000 crore) has been calculated and is approved by other Non-Banking Financial Companies (NBFCs) and Commercial Banks. Figure 4 shows the annual investment roadmap in renewable energy with a prime focus on solar energy power. It is quite apparent that investment is progressively increased from Rs.58,000 crores to Rs.102,000 crores from the year 2015–2016 to 2016–2017 owing to various government funding and debt and equity based funding's form various institutions. Also, it is anticipated to be raised to Rs.149,000 crores at the end of year 2021.

Furthermore, to transmit the clean energy from the sites producing the renewable energy to the sites consuming it, a policy named green energy corridor—consisting of seamless transmission lines for smooth flow of the dedicated renewable electricity has been proposed by Minister of State for Power, Coal, New and Renewable Energy and Mines, Government of India. Interstate and intrastate corridors of transmission lines are expected to cost around 80 billion and 45.75 billion rupees, respectively. To help this, the German Development Bank is giving the loan of about 1 billion euros. Also, 'Make in India' Campaign of current Prime Minister Narender Modi has made various government institutions to discuss more intensely about innovative financial methods to reduce cost of debt such as green bonds of infrastructure debt funds (IDF), dollar denominated tariffs, Hedging risk funds to allow cheaper financing. Innovative financing mechanisms also include [13]:

- Renewables accorded priority sector lending status—loans up to 15 crore for renewable power projects included. Banks mandated to disburse 40% of adjusted net bank credit to priority segments.
- Rooftop solar through home loans—DFS has issued guidelines and circulars issued by various banks to include rooftop under housing loan.
- Tax-free bonds—government approves INR 50 billion tax-free bonds to enable low interest rate funding for renewable energy.
- Green Bonds: Enabling low cost funding for clean energy, REC, PFC, IREDA, IDBI and private sector like ICICI, Yes Bank asked to issue green bonds to raise funding for RE projects.

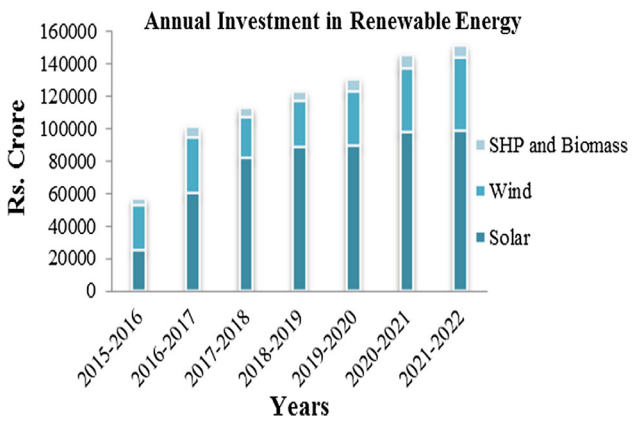


Fig. 4 Road map of investment in solar projects [13]

Future Solar Projects in India

India having high solar insolation is very ideal place for solar electricity generation through CSP or PV. Quite recently, a 35,000 km^2 area of Thar dessert is proposed for large solar power projects, to generate around 2100 GW of power. Today, among 38,283.59 MW of total installed capacity around 4684.74 MW of electricity is generated by solar projects which is around 12% of the total installed capacity. As a matter of fact, India is expected to install an additional 10,000 MW by 2017 and a total of 1, 00,000 MW by 2022. Currently, the French group AREVA solar is constructing and installing a 250 MW CSP plant that has actually grounded a milestone in Rajasthan being largest CSP installation in Asia this far. Breathtakingly, Mega Green Solar Power Projects is also planning and installing a whopping 4000 MW solar project near Sambhar Lake in Rajasthan. Once it is completed, it would be biggest solar project on the planet.

Conclusion

The central gist of this work could be most significantly digressed towards the application of solar technology which actually offers a diverse variety of application. The solar power technologies like PV and CSP have been discussed with focus on their techno-economic constraints of implementation. PV and CSP both

technologies with the perspective of India were considered in view of JNNSM. The JNNSM has some serious ambitious goals of producing 20 GW of solar electricity by 2020 which was later increased to 100 GW. The advancement and achievement of the mission was discussed in three phases of the mission. Even though JNNSM is a great breakthrough towards the renewable energy generation and its involvement in the grid, the execution part lacked a little consistency. As mentioned in the earlier—power production through CSP was subjected to a 30% of reduction, also a planned 1500 MW CSP project was changed to 750 MW PV. Suggestion with regards to this—implementing some advanced project development techniques so as the proposed projects could be finished on time, can be very much appreciated. The Government of India's Solar Mission is a visionary and inspiring policy measure that has the potential to be a leading example for the world. Harnessing the potential of solar energy for the benefit of rural areas of India will enlighten the hearts and minds of the rural population. The advancements and encouragements of the Solar Energy technology carried out so far can be summed up as follows:

- Amendments in electricity supply and regulatory acts to make sure that a minimum of solar electricity is generated and transmitted to the grids within the purchase plan.
- Measures to ease the minimum interest loan with soft repayment and other financial support to manufactures and users to promote rapid commercialisation.
- Reduction of customs, sale tax exemption and subsidised renewable power pertaining to solar energy.
- Awareness generation and capacity building.
- Market generation and support such as supply and service networks.

References

1. District Census Hand Book (2011) http://www.censusindia.gov.in/2011census/dchb/DCHB.html
2. Deendayal Upadhyaya Gram Jyoti Yojana (DDUGJY) (2011) http://powermin.nic.in/en/content/deendayal-upadhyaya-gram-jyoti-yojana-ddugjy
3. G. Garud, Rural electrification: challenges and the way ahead (2015), http://powerwatchindia.com/ruralelectrification-challenges-and-the-way-ahead/
4. J.P. Kesari, R.S. Mishra, A. Uppal, M. Arora, Skill development and empowerment of rural India through renewable energy education, training and innovation: the role of Delhi Technological University', in *International Conference of Advance Research and Innovation (ICARI-2015)* (2015), pp. 41–42
5. India Solar Resource Maps & Data (2016), http://webcache.googleusercontent.com/search?q=cache:http://www.nrel.gov/international/ra_india.html&gws_rd=cr&ei=NCawV6evH8H8vgSVqr34Bg
6. India's Ambitious Renewable Energy Plans (2015) http://www.gwec.net/indias-ambitious-renewable-energy-plans/
7. Scheme/Documents (2013), http://www.mnre.gov.in/solarmission/jnnsm/introduction-2/
8. C. Bhushan, The state of concentrated power in India, (2015), www.cseindia.org/userfiles/State%20of%20CSP%20in%20India%20low.pdf
9. Grid Connected PV (2016), http://www.seci.gov.in/content/innerpage/grid-connected-pv.php
10. J.P. Kesari, U.K. Sahoo, S.K. Singh, P.B. Sharma, Transforming India into a green nation: bottom up approach through solar energy education, training and innovations. Int. J. Sci. Res. **3**(6), 153–156 (2014)
11. Y.P. Chandra, A. Singh, S.K. Mohapatra, J.P. Kesari, L. Rana, Numerical optimization and convective thermal loss analysis of improved solar parabolic trough collector receiver system with one sided thermal insulation. Sol. Energy **148**, 36–48 (2017)
12. J. Joy, D. Joy, T.S. Panwar, People's Perception study—Renewable Energy in India 2014, WWF-India (2015) http://awsassets.wwfindia.org/downloads/peoples_perception_study__renewable_energy_in_india_2014.pdf
13. K.S. Popali, Current status, outlook and investment opportunities in RE sector in India, IREDA, New Delhi (2016) http://www.eclareon.com/sites/default/files/Presentations/04_kuljit_popli_-_ireda_-finanzierung_von_solarprojekten.pdf

Volume 1, April 2018

"Solar Thermal Technology: Potential and Industrial Applications"

Piyush Rawat
Department of Mechanical Engineering, M.Tech,
Delhi Technological University

J. P. Kesari
Associate Professor, Department of Mechanical Engineering,
Delhi Technological University,

Abstract

India's electricity sector has an installed capacity of 334.4 GW as on Jan 31st 2018, with a renewable share of only 18.37%. By 2022, Indian government have targeted 175GW of energy from renewable sources with 100 GW from solar power. Solar thermal technology uses the sun's energy, rather than fossil fuels, to generate low-cost, environmental friendly thermal energy. This energy is used to heat water or other fluids, and can also power solar cooling systems. This paper discusses about the potential of solar thermal in India and industrial utilization.

1. INTRODUCTION

Solar energy is a renewable and abundant source of energy. It is also a sustainable solution for industrial processes where proper implementation will drive the industries toward zero carbon emission in future. Major industrial processes dependence on either fossil fuel or electricity to supply heat. Depletion of non-renewable sources and high greenhouse gas emission (GHG), a shift toward renewable energy is essential. The real disadvantage which hinders the scope for small and medium-sized industries to integrate renewable energy system is high installation cost and mostly to ensure a proper and continuous supply of energy.

Solar thermal systems in industrial processes and electricity production are becoming popular in different parts of the world but heavily dependence on radiation intensity. The storage system can be considered as supplementary, depending on seasonal variation, i.e. reduction of solar intensity during winters. It may improve the overall performance of the system.

2. THE SOLAR THERMAL POTENTIAL IN INDIA

Energy gained from solar irradiation after heat conversion is called solar thermal energy. Solar energy is an important renewable energy which will play a vital role in the future for supplying energy [1-4]. During the past several decades, concentrated solar power (CSP) technology became an important option for harnessing solar energy [5-8]. In January 2010, Government of India started the Jawaharlal Nehru National Solar Mission (JNNSM), and activities toward solar-based power generation took a considerable momentum [9]. The

mission aims for the deployment of 20,000 MW of solar power by the year 2022 and creating a policy framework for it [10].

The estimated solar power generation for a threshold direct normal irradiance (DNI) value of 1800 kWh/m^2 is 756 GW and for a threshold DNI value of 2000 kWh/m^2 is 229GW [11] (with all the waste-lands having wind speed of 4m/s or more allocated for wind power production). Hence a large area of waste-land in India can be utilized for production of wind and solar energy.

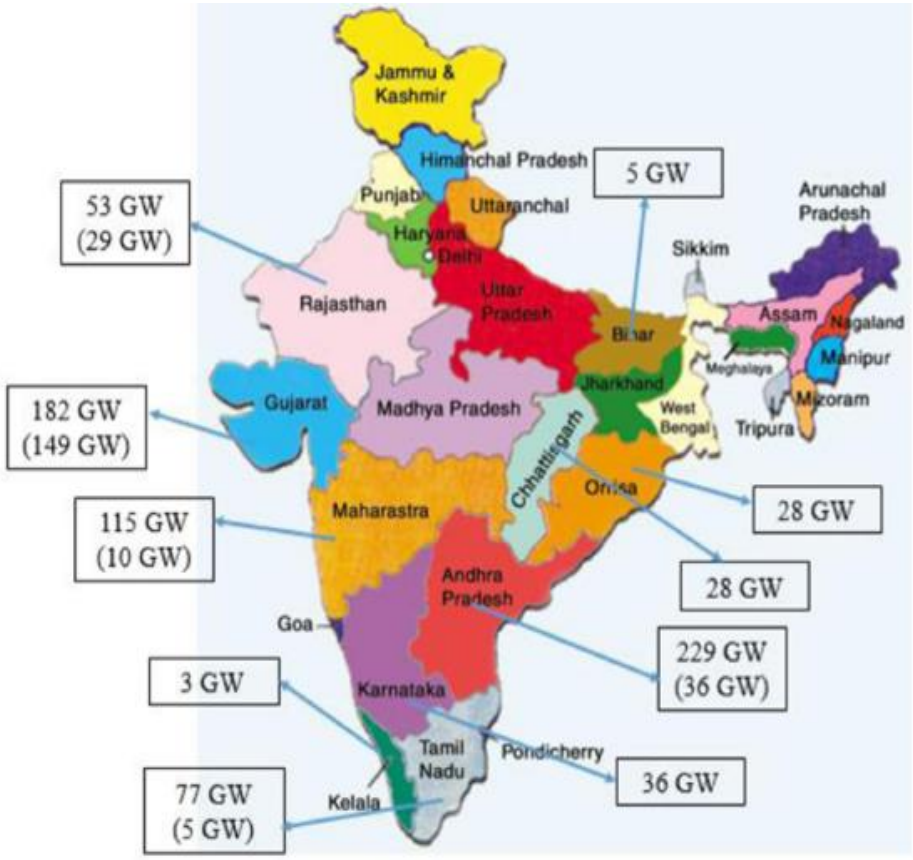


Figure 1 : CSP potential in different states. (Annual threshold of 1800 kW/m^2; Data in the bracket denotes CSP potential for DNI of 2000 kWh/m^2) [11]

3. HARNESSING SOLAR ENERGY FOR INDUSTRIAL PROCESS HEATING

The various energy sources used and amount of energy shared by individual sources, consumed by the industries vary across regions and countries, technological choice, and availability of fuels. According to a report of International Energy Agency (IEA), most of the energy demands globally are met with fossil fuels [15]. With the rapid increase in energy demand, the rate of fossil fuel consumption will increase**,** causing high GHG emission and global warming. Hence, it is a desire to reduce fossil fuel consumption while maintaining the increased energy demand. It can be achieved by (i) improving fuel efficiency in various

processes and equipment [12-14], (ii) replacing fossil fuels with solar thermal energy for process heat applications.

Conversion of solar energy into thermal energy has comparatively higher efficiency than producing electricity from solar energy. For example, the efficiency of producing electricity from solar energy is between 15-20% while for heat production, the corresponding values are between 20-25% [16-18]. Also, a temperature range of 50-300^{o}C is required for major part of the industrial process [18] and can be achieved with proper technology.

Table 1: Various industrial processes and heating temperature.

Industry	**Process (es)**	**Temperature (^{o}C)**	**Medium**
Food processing, milk processing and beverages production	Cooking Pasteurization Sterilization Tempering Drying Washing Heat treatment	70-120 60-150 100-140 40-80 40-100 40-80 60-80	Steam, water Steam Steam Steam Air Steam, Air Water
Textile	Dying Drying Pressing Fixing Printing	60-90 100-130 120-140 160-180 40-130	Water Steam Steam Steam Water, steam
Pulp and Paper	Bleaching De-linking Drying Pulp preparation Boiler feed water heating	120-150 60-90 90-200 120-170 60-90	Water Steam Air, steam Pressurized water Water
Chemical and pharmaceutical	Distillation Evaporation Drying Thickening	100-200 110-170 120-170 130-140	Water Steam Air, Steam Steam
Automobile	Paint pre- treatment Baking of paints Paint drying	40-50 175-225 160-175	Water Steam Air
Steal making and other casting industries	Hardening, annealing, tempering, forging, rolling	700-1500	Air
Cement manufacturing	Lime calcing	600-1200	Air

3.1 Heat transferring media

Heat transferring media used in industries for process heating depends on the requirement of specific application or process and it may or may not have direct contact with the other

working fluids. The important characteristics of heat transfer media for process heating are: (i) low vapour pressure, (ii) high heat capacity, (iii) high thermal stability and (v) low corrosiveness [19-21].

Steam is the most common heat transferring media used in industries for process heating. For e.g., in the USA, around 37% of fossil fuel is burned for the production of steam [22] and up to one-third of all energy consumption is used for steam generation in the UK [23]. The heat content of steam is mostly stored as latent heat and a large quantity of heat can be extracted efficiently at a constant temperature. Also due to high energy density than other media like hot water and thermal oil, it is best suitable for process heating.

3.2 Solar collectors

A solar collector is a device which absorbs solar radiation from the Sun, and the extracted heat is transferred to the working fluid or heat transfer media. Selection of appropriate solar collector mainly depends on four factors (i) operation temperature, (ii) efficiency of solar collector, (iii) cost and (iv) energy received annually [24-28]. Factors like availability of space and possibility of root top installation are also important.

Mainly three types of solar collector technologies are used for Solar Industrial Process Heat (SIPH) applications – flat plate, concentrated, and evacuated tube. For flat plate collectors, both water and air is used as heat transfer media. For low-temperature industrial processes, flat plate and evacuated tube collector based system are used [29]. Concentrated type collector can produce a heating temperature of up to 300^{o}C and most commonly used concentrated collectors are the parabolic trough, paraboloid dish and linear Fresnel reflector [30-32]. Table 2 gives the detail about different solar collectors.

Table 2: Type of solar collector, tracking type, temperature range and efficiency.

Solar Collector	Tracking	Temperature Range (^{o}C)	Efficiency
Flat Plat	Non tracking	30-100	$\eta i=0.732–3.411*(Tm-Ta/I–0.294*(Ti-Ta)^2/I$ [33]
Evacuated tube	Non tracking	50-100	$\eta i=0.62–0.395*(Tm-Ta/I-0.002*(Ti-Ta)^2/I$ [34]
Parabolic trough collector (PTC)	Single axis	60-300	$\eta i=0.71–0.3581*(Tm-Ta/I-0.0019*(Tm-Ta)^2/I$ [35]
Arun Dish 160	Double axis	Up to 250	$\eta i = 0.65–0.40*(Tm-Ta)/I$ [36]
Linear Fresnel reflector	Single axis	Up to 300	$\eta i=0.635–0.0265*(Tm-Ta/I-0.00043*(Tm-Ta)^2/I$ [37]

3.3 SIPH systems: Design and integration

Solar industrial process heating (SIPH) system basically consists of collectors, heat transfer media, heat exchangers, thermal storage, and auxiliary backup as an optional component. Figure 2 shows a schematic diagram of solar concentrated based process heating system for industries. Solar irradiations falling on the thermal collector is absorbed and then transferred to the working fluid, generally water. Heat rejected to working fluid is then utilized for various processes based on the industries.

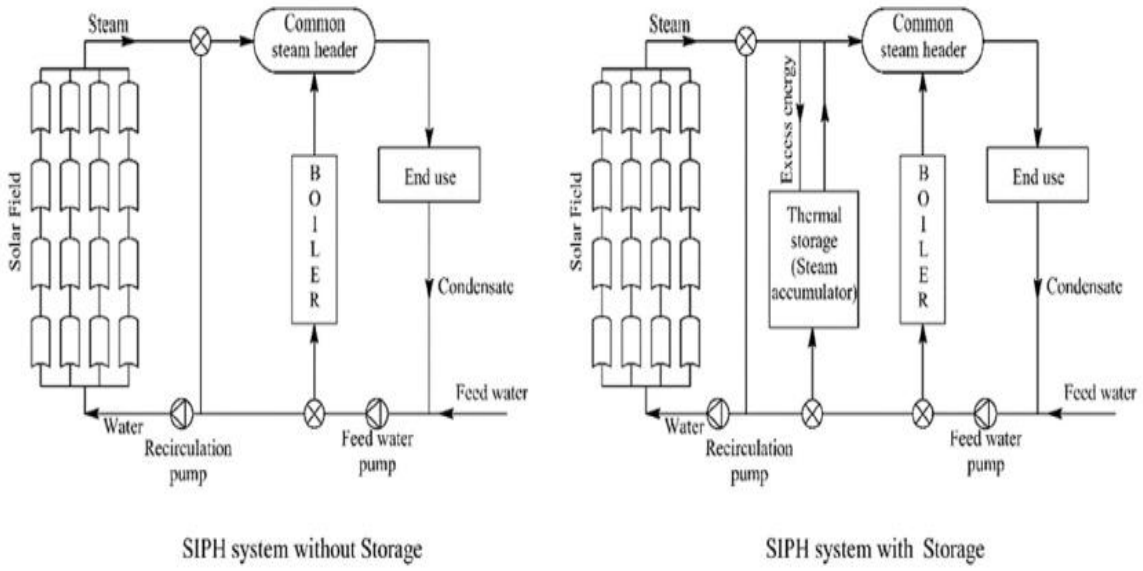


Figure 2: Schematic diagram of SIPH systems [9].

Indirect steam generation

Water/steam is not the working fluid, instead thermal oil, glycol or PCM is heated by the solar collectors and is then used to generate steam indirectly using heat exchangers. It is further configured into – flash boiler based and unfired boiler based systems. In flash boiler system, pressurized water is heated by the solar collectors and then passed through a throttling valve, due to which a part of it, is converted to steam. This steam is delivered and utilized for industrial applications. And the part of water is re-circulated back to the collectors. Feed water tank and circulation pump is also required. In the unfired boiler based system, working fluid passes through the collectors where heat is absorbed, and is then rejected to secondary fluid like air and water.

Direct steam generation

In this mode, water is the working fluid. It is boiled by the solar collectors and the generated steam is stored under high pressure with thick insulation or utilized directly. Steam

is separated from the water in a steam drum. The water is then re-circulated back to the solar field. Feed water tank is also provided.

In industries, a continuous supply of process heat is demanded, hence hybrid SIPH systems are accepted. An auxiliary boiler is used as a backup during cloudy hours and night. A hybrid system is shown in fig. 3.

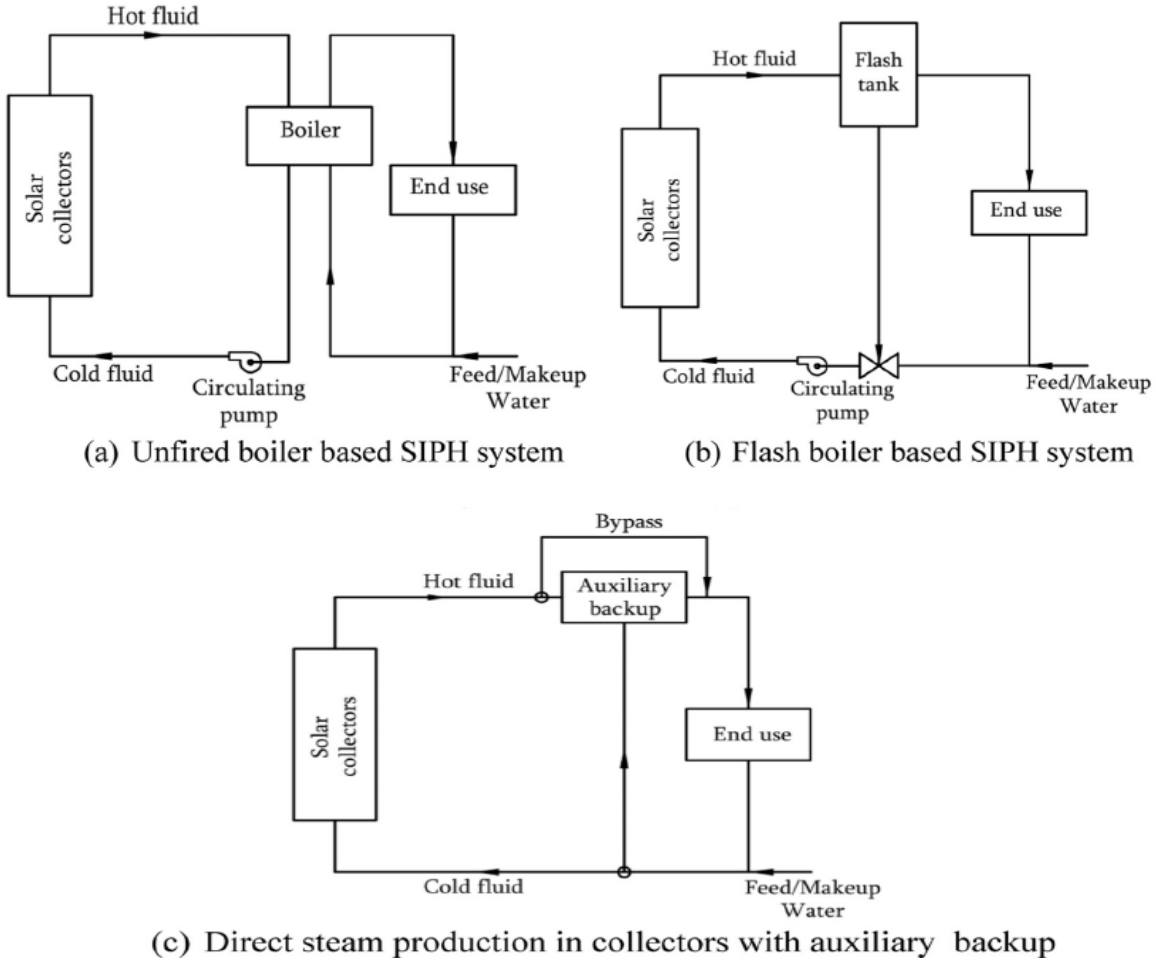


Figure 3: SIPH systems [9].

Solar collector area

Calculation of solar collector area depends on (i) DNI or GHI (Global horizontal irradiance) values (DNI in case of concentrated based system and GHI in case of stationary collectors) (ii) process heating demand (iii) efficiency of collectors.

$$SCA = \frac{PHR_h}{(\eta_c)\ (GHI_d)}$$

SCA = solar collector area; PHR_h = rate of process heat requirement; η_c = thermal efficiency of solar collector; GHI_d = Global horizontal irradiance (W/m^2)

4. PROBLEMS ASSOCIATED WITH LARGE SCALE DEPLOYMENT OF SIPH SYSTEM

Potential of solar thermal energy is very high but harnessing it is quite difficult. Various problems associated are discussed in the literature work. Several reasons are taken from them [38-39]. One of the significant barriers to the installation of the solar thermal process is the structure of the industrial sectors. Integration of solar thermal system into existing large-scale plants requires large space, heavy optimization and thorough knowledge of the system. All solar thermal process system requires a heavy initial cost. For small and medium scale industries, due to lack of financial support, the main barrier is the upfront cost and the total installation cost, even if the overall lifetime cost would be low. Also in many countries government provides subsidy or discount on fossil fuels to support industrial development. Hence small and medium scale industries prefer continuous heat from fossil fuels.

Lack of skilled labours and public awareness is also an important barrier for deployment of solar process heat system. Many small and medium scale industries do not have enough information or first-hand experience with solar thermal technology. They also prefer conventional sources which can provide a continuous supply of heat since they cannot deal with initial loss due to inconsistent heat supplied by solar technologies. Hence solar thermal technologies are being avoided by most of the industries.

There are only few research institutes with thorough and proper knowledge of solar thermal energy. To overcome this problem training and distribution of existing knowledge and concepts are needed.

CONCLUSION

For several industrial processes as shown in table 1, the temperature range required can be achieved with the proper installation of solar thermal systems. Adaption of solar thermal technologies for small and medium scale industries is beneficial for the growth of the industries as well as a sustainable development of the country. Solar collectors like the flat plate and evacuated tube can provide an operating temperature of about 30-100^{o}C. For the high temperature range of about 250-300^{o}C, parabolic trough, and linear Fresnel reflector are used in industries.

There is a huge potential of adopting SIPH systems in industries for process heating. But problems like installation of new systems into systems which are already running at full potential, installation cost, the area required for implementation, storage problem etc. are some factors which may inherit. Hence adaption of proper policies and regulations are necessary.

REFERENCES

[1] Tsoutsos T, GekasV, Marketaki K.Technical and economical evaluation of solar thermal power generation. Renew Energy 2003;28(6):873–86.

[2] Brakmann G, Aringhoff R, Geyer M, Teske S. Concentrated solar thermal power – now. Amsterdam: Green peace International; 2005.

[3] McKinsey. The economics of solar power. The McKinsey Quarterly, June2008.

[4] European Commission. Concentrating solar power – from research to implementation. Luxembourg: European Commission; 2007.

[5] Yogev A, Kribus A, Epstein M, Gogan A.Solar tower reflector systems : a new approach for high-temperature solar plants. Int J Hydrog Energy 1998;23(4):239–45.

[6] Segal A, Epstein M. Comparative performances of tower-top and tower-reflector central solar receivers. Sol Energy 1999;65(4):207–26.

[7] Kodama T. High-temperature solar chemistry for converting solar heat to chemical fuels. Prog Energy Combust Sci 2003;29(6):567–97.

[8] IEA. Technology roadmap – concentrating solar power. Paris, Cedex, France: International Energy Agency; 2010.

[9] Sharma A. A comprehensive study of solar power in India and World. Renew Sustain Energy Rev 2014;15(4):1767–76.

[10] MNRE. Jawaharlal Nehru National Solar Mission (JNNSM) – towards building solar India. (See). New Delhi: Ministry of New and Renewable Energy, Government of India; 2010([accessed 25 january 2018.]).

[11] Chandan Sharma, Ashish K. Sharma, Subhash C. Mullick, Tara C. Kandpal, Assessment of solar thermal power generation potential in India, Renewable and Sustainable Energy Reviews, Volume 42, 2015, Pages 902-912, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2014.10.059.

[12] Abdelaziz EA, Saidur R, Mekhilef S. A review on energy saving strategies in industrial sector. Renew Sustain Energy Rev 2011;15:150–68.

[13] Saidur R, Rahim NA, Ping HW, Jahirul MI, Mekhilef S, Masjuki HH. Energy and emission analysis for industrial motors in Malaysia. Energy Policy 2009;37:3650–8.

[14] Saidur R. A review on electrical motors energy use and energy savings. Renew Sustain Energy Rev 2010;14:877–98.

[15] US Energy Information Administration (EIA). International Energy Statistics Database See: < ⟨http://www.eia.gov/oiaf/aeo/tablebrowser/#release/Table⟩ > [accessed 27 January 2018]; 2015.

[16] Chu Y. Review and comparison of different solar energy technologies. Global Energy Network Institute; 2011, ⟨http://geni15.wrsc.org/sites/default/files/Review-and-Comparison-of-Different-Solar-Technologies.pdf⟩, [accessed 28 February 2018].

[17] Byrne J. World solar energy review: technology market and polices. Centre of Energy and Environment Policy University of Delaware; 2010, ⟨http://ceep.udel.edu/wp-

content/uploads/2013/08/2010_World_Solar_Energy_Review_Technology_Markets_and_Policies1.pdf⟩, [accessed 28 January 2018].

[18] Vejan K, Lauterbach C, Schmitt C. Solar heat for industrial processes –Potential, technologies and applications. In: Proceedings of the International Conference on Solar energy for MENA region (INCOSOL); 2012.

[19] Benoit H, Spreafico L, Gauthier D, Flagman G. Review of heat transfer fluids in tube-receivers used in concentrating solar thermal systems: properties and heat transfer coefficients. Renew Sustain Energy Rev 2016;55:298–315.

[20] Vignaroobana K, Xinhai B, Arvayc A, Hsua K, Kannana AM. Heat transfer fluids for concentrating solar power systems – A review. Appl Energy 2015;146:383–96.

[21] Singh YK. Heat transfer fluids and systems for process and energy applications. CRC Press; 1985, [ISBN 9780824771911].

[22] Einstein D, Worrell E, Khrushch M. Steam systems in industry: energy use and energy efficiency improvement potentials, Lawrence Berkeley National Laboratory 2012. See: ⟨http://osti.gov/bridge/servlets/purl/789187-uTGqsP/native/⟩.

[23] Carbon Trust. Process Heating Introducing Energy Saving Measure for Business, Private Sector Energy Efficiency. UK Department for International Development (DFID) (2015).

[24] Rabl A. Yearly average performance of the principal solar collector types. Sol Energy 1981;27:215–33.

[25] Adstena M, Perersb B, Wackelgarda E. The influence of climate and location on collector performance. Renew Energy 2002;25:499–509.

[26] Andersen E, Furbo S. Theoretical variations of the thermal performance of different solar collectors and solar combi-systems as function of the varying yearly weather conditions in Denmark. Sol Energy 2009;83:552–65.

[27] Horta P, Brunnerb C, Kramera K, Frank E. IEA/SHC T49 Activities on process heat collectors: available technologies, technical-economic comparison tools, operation and standardization recommendations. Energy Procedia 2016;91:630–7.

[28] Sardeshpande V, Pillai I. Effect of micro-level and macro-level factors on adoption potential of solar concentrators for medium temperature thermal applications. Energy Sustain Dev 2012;16:216-23.

[29] Colangelo G, Favale E, Miglietta P, Risi A. Innovation in flat solar thermal collectors: a review of the last ten years experimental results. Renew Sustain Energy Rev 2016;57:1141–59.

[30] Kalogirou SA. Chapter 3 – Solar energy collectors [Second Edition]. Sol Energy Eng 2014:125–220.

[31] Kalogirou SA. Solar thermal collectors and applications. Progress Energy Combust Sci 2004;30(3):231–95.

[32] Sharma V Rakesh, Bhosale SJ, Kedare SB, Nayak JK. Field tests of the Performance of Paraboloid Solar Concentrator-ARUN160 at Latur, In: Proceedings of the National Conference on Advances in Energy Research, Indian Institute of Technology –Bombay, December 4-5, In: Proceedings of the NCAER;182-187; 2006.

[33] SRCC. Flat plate collector. Certification number, 2011129D. Solar Rating and Certification Corporation. See at: < ⟨https://secure.solar-rating.org/Certification/Ratings/RatingsSummaryPage.aspx?Type=1⟩ > [accessed 28 January 2018];2013.

[34] SRCC. Evacuated tube collector. Certification number, 10001929. Solar Rating and Certification Corporation. See at: < ⟨https://secure.solar-rating.org/Certification/Ratings/RatingsSummaryPage.aspx?Type=1⟩ > [accessed 28 January 2018];2014.

[35] SRCC. Parabolic trough collector. Certification number, 10001962 Solar Rating and Certification Corporation. See at ⟨https://secure.solar-rating.org/Certification/Ratings/RatingsSummaryPage.aspx?Type=1⟩ [accessed 30 January 2018]; 2014.

[36] Bhosale SJ, Kedare SB, Nayak JK. Performance of ARUN 160 concentrating solar collector installed at latur for milk pasteurization. SESI J 2008;18:23–31.

[37] Horta P. Process Heat Collectors: State of the Art and available medium temperature collectors. Solar Process Heat for Production and Advanced Applications. IEA SHC Task 49. See: < ⟨http://task49.iea-shc.org/⟩ > [accessed 30 January 2018]; 2015.

[38] European Solar Thermal Industry Federation. See:<⟨http://www.estif.org/fileadmin/estif/content/policies/downloads/D23-solar-industrial-process-heat.pdf⟩ > [accessed 30 January 2018].

[39] IRENA. Solar Heat for Industrial Processes Technology Brief. International Renewable Energy Agency. See: < ⟨http://www.solarthermalworld.org/sites/gstec/files/news/file/2015-02-27/irena-solar-heat-for-industrial-processes_2015.pdf⟩ > [accessed 30 January 2018]; 2015.

1st International Conference on Recent Trends and Innovation in Engineering, Sciences and Management
Adesh Institute of Engineering & Technology , Faridkot ICRTIESM-18
14th-15th April 2018 www.conferenceworld.in ISBN: 978-93-87793-24-8

Status of Ambient Air Quality in Khondbond Iron Ore Mines, Orissa (India)

Ankita Srivastava[1], Amit Krishan[2*], Rajeev Kumar Mishra[2]

[1]*Department of Environmental Science and Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad, Jharkhand, India*

[2]*Department of Environmental Engineering, Delhi Technological University, Delhi, India*

ABSTRACT

Khondbond Iron Mine is located at Keonjhar district of Orissa, around 18 km from Joda East Iron Mine. Iron ore available in this mine is especially suitable for sponge iron making. Mining is a vital industry for the industrial and economic growth of any country and is inherently disruptive to the environment. Due to unsuitable and inefficient working practices during mining causes environmental deterioration. The current work has been carried out to assess the ambient air quality in the open cast iron ore mines in the Khondbond of Orissa, India. For the decision makers, it has been an important task to keep the air quality within an acceptable limit. In the present study monitoring and analysis of SO_2, NO_2, PM_{10} and $PM_{2.5}$ have been carried out. The Air Quality Index (AQI) was also calculated using IND-AQI specified by Central Pollution Control Board (CPCB), New Delhi and it was based on the dose-response relationship of various pollutants. The result illustrates that all the four parameters were well below the maximum allowed limit of National Ambient Air Quality Standards (NAAQS) and the particulate materials were the major pollutants in mining areas for different areas. The overall AQI was found to fall under the category of good.

Keywords: *AQI, Iron ore mine, Mining, NAAQS, Open cast mine*

1. Introduction

The air we breathe is a mixture of gases and small solid and liquid particles. Some substances come from natural sources while others are caused by human activities. Keeping the air quality acceptable has become an important task for decision makers as well as for non-governmental organizations [2]. Air pollution occurs when the air contains substances in quantities that could harm the comfort or health of humans and animals, or could damage plants and materials. These substances are called air pollutants and can be either particles, liquids or gaseous in nature [1]. Mining is the extraction of valuable minerals or other geological materials from the earth, usually from an ore body, vein or (coal) seam. Minerals play an important role in the economic development of the country as minerals are the basic raw materials to promote the growth.

Mining is a vital industry for the industrial and economic growth of any country. The mineral industry in India is reckoned not only as an important contributor to the country's GDP and foreign trade, it is also one of the major industries that absorb a considerable amount of the country's working population. The Indian mining industry provides employment to about 1.1 million persons. India's mining industry is projected to touch over $30 billion (about Rs. 1, 27,662 crores) accounting for about 2.5% of the GDP in the next four years, the latest report said.

Mining activities are inherently disruptive to the environment. The environmental deterioration caused by mining activities occurs mainly as a result of inappropriate and wasteful working practices. Over the past few years, with the introduction of mechanized mining techniques and heavy earth moving equipment, the problem has been further aggravated [3] [4]. The magnitude of impact on air quality depends upon the types of mineral being mined, methods, scale and concentration of mining activities [5]. Open-cast mining is more deteriorating to air quality than underground mining. Land clearing, removal of overburden, and vehicular movement on the haul roads, excavation, loading and unloading of ore materials as well as overburden are the major sources of atmospheric emissions from open-cast mining activity.

Air quality monitoring and assessment are required to prevent and minimize the deterioration of air quality due to mining and other anthropogenic activities. The existing evaluation and assessment of air quality in India are based entirely on the compliance of measured concentration levels of pollutants with the National Ambient Air Quality Standards (NAAQS) [6]. This approach is used to maintain the desired quality of the ambient atmosphere, the concentration level of pollutants found below the NAAQS are fair and acceptable air quality, but sometimes the concentrations are sufficiently high to pose serious environmental and health problems.

The use of standards is important in administrating and enforcing the desired policy, but they are not a complete tool for evaluating the environmental quality [6]. The need to quantify air quality and follow the evolution of pollution has led to a large amount of Air Quality Indices, giving an idea of the state of pollution in a day [7]. Air Quality Indices are also of specific significance because; air pollution monitoring data are generally complex and not understandable to the general public. The large data often do not convey the air quality status to the scientific community, policymakers and the general public in a simple and straightforward manner. This problem can be addressed by determining the Air Quality Index (AQI) of a given area using environmental synthetic indices to summarize complex situations in a single figure, allowing for comparisons in time and in space [8].

The current work has been carried with the objectives of assessment of the ambient air quality in the open cast iron ore mines in the Khondbond of Orissa, India.

2. Materials and methods

2.1. Study Area

Khondbond Iron Mine is located at Keonjhar district of Orissa, around 18 km from Joda East Iron Mine. The Khondbond mine has deposits of sponge grade iron ore and also occurrences of manganese ore. The iron ore

deposit in this mine is worked by mechanised open cast mining in a series of 6 m high benches. The Crushing Plant at Khondbond, established in 1994, crushes the ore in three stages to produce lump ore, which is dispatched for sponge iron making to M/s. Tata Sponge Iron Ltd. (TSIL) at Joda, while the fines are transported to the Steel Works at Jamshedpur.

Discovered in 1917, the integrated manganese complex has the capacity to handle 0.3 Million MT of ore and 40000 tones per annum of manganese alloys. The plant recorded its highest ever production and crossed breakeven production with lowest ever cost. Iron ore available in this mine is especially suitable for sponge iron making. The crushing plant established in 1994 by Svedala, crushes the ore in three stages to produce lump ore with a size range of 18 mm to 5 mm. It is then dispatched for sponge iron making to M/s Tata Sponge Iron Ltd at Joda, while the fines are transported to Jamshedpur for steelworks.

2.2. Monitoring and Analysis

Total four monitoring sites were selected for the present study. The first sampling site was near the plant (A–1). The second sampling site was at the mining site (A–2). The third sampling site was near weighbridge (A–3). The fourth sampling site was at the site of Equipment Maintenance (A–4).

The samples of SO_2, NO_2, PM_{10} and $PM_{2.5}$ were collected at all the four sites. SO_2, NO_2, PM_{10} and $PM_{2.5}$ samples were collected for 24 hr at four different sites for 2 days per week for four weeks. The concentration of NO_2 was measured with the standard method of Modified Jacobs- Hochheiser method (1958), SO_2 was measured by Modified West and Geake method (1956). SO_2 and NO_2 samples were collected by aspirating air through measured volume (25 ml) of 0.04 M potassium tetra chloro mercurate (K_2HgCl_4: TCM) and a mixture of sodium hydroxide and sodium arsenite ($NaOH + Na_2AsO_2$), respectively at a flow rate of 2 l/min for 24 hr. The collected samples were transferred to pre-washed polypropylene bottles, and preserved in the refrigerator and analysed colorimetrically at 548 nm for SO_2 and 540 nm for NO_2. The apparatus was kept at a height of 2 m from the surface of the ground.

PM_{10} and $PM_{2.5}$ samples were collected on Whatman GF/A and Teflon (Millipore) filter papers by Respirable Dust Sampler (APM 460DX, Envirotech, New Delhi) and Wins-Anderson impactor (APM 550, Envirotech, New Delhi) with the sharper cut point of 10 µm and 2.5 µm, respectively. The Respirable Dust Sampler and Wins-Anderson impactor were operated at flow rates of 1.0 m^3/min and 16.671 m^3/min (or 1 m^3/hr), respectively.

All the filter papers were pre-weighed on a Metler analytical weighing balance before the sampling and desiccated for 24 hr. To avoid the contamination, the conditioned and weighed filter papers were placed in filter holder cassette ($PM_{2.5}$) and zip lock polybag (PM_{10}) and were taken to the field for sampling. Before loading the filter papers on the samplers, initial volume and timer readings were noted for $PM_{2.5}$ and the manometer reading for PM_{10} sampler. Then the filter papers were loaded on respective samplers and screwed properly before starting the samplers. After sampling, the loaded filter of $PM_{2.5}$ was removed with forceps and placed in a cassette and wrapped with aluminium foil. Similarly, the PM_{10} filter paper was wrapped in aluminum foil and

placed back in zip lock polybag and both the filter papers were brought back to the laboratory. In the laboratory, filter papers were conditioned and weighed again to determine the mass concentration of the PM_{10} and $PM_{2.5}$. The weighed filter papers were preserved in the freezer until further chemical analysis.

2.3. Air Quality Index (AQI)

Air Quality Index (AQI) is a tool, introduced by Environmental Protection Agency (EPA) in the USA to measure the levels of pollution due to major air pollutants. An AQI is defined as an overall scheme that transforms weighted values of individual air pollution related parameters into a single number or set of numbers [9]. In the present study the AQI was calculated using IND-AQI specified by CPCB. The index has been developed based on the dose-response relationship of various pollutants. AQI concept transforms weighted values of individual air pollutants into a single number or set of numbers which may be widely used for air quality communication and decision making. This IND-AQI has 6 categories.

The AQI method involves the formation of sub-indices for each pollutant and aggregation of sub-indices. It has been developed on the dose-response relationship of various pollutants [10]. Table 1 shows the Linear segmented relationship for sub-index values and the corresponding pollutant concentrations that are calibrated to Indian conditions.

TABLE 1. Various Categories of IND-AQI and Break Points of Various Pollutants (National Air Quality Index, CPCB, October 2014) (Units: µg/m^3)

AQI Category	Range	PM_{10} (24hr)	$PM_{2.5}$ (24hr)	SO_2 (24hr)	NO_2 (24hr)
Good	0-50	0-50	0-30	0-40	0-40
Satisfactory	51-100	51-100	31-60	41-80	41-80
Moderately Polluted	101-200	101-250	61-90	81-380	81-180
Poor	201-300	251-350	91-120	381-800	181-280
Very Poor	301-400	351-430	121-250	801-1600	281-400
Severe	401-500	>430	>250	>1600	>400

The mathematical equations for calculating sub-indices is as follows

$$I_P = \left(\frac{I_{HI} - I_{LO}}{B_{PHI} - B_{PLO}} X \left(C_P - B_{PLO} \right) \right) + I_{LO}$$

Where

I_P is AQI for pollutant "P" (Rounded to the nearest integer),

C_P the actual ambient concentration of pollutant "P",

B_{PHI} the upper-end breakpoint concentration that is greater than or equal to C_P,

B_{PLO} the lower end breakpoint concentration that is less than or equal to C_P,

I_{LO} the sub-index or AQI value corresponding to B_{PLO},

I_{HI} the sub-index or AQI value corresponding to B_{PHI}

3. Results and discussion

To assess the air quality status of the study area, the CPCB guidelines for sampling and measurement of ambient air quality parameters were followed. The NAAQS [11] for $PM_{2.5}$, PM_{10}, SO_2 and NO_2, is depicted in Table 2. The concentration levels of $PM_{2.5}$, PM_{10}, SO_2 and NO_2 in the Khondbond iron ore mine are presented in Table 2.

3.1. Near plant (A–1)

The sampling point A–1 was near crushing plant. The main source of pollution near the plant is vehicular pollution, crushing, screening and processing. The concentration of SO_2 at near crushing plant varies from 9.4 to 9.90 µg/m^3, which is within the permissible limit of 80µg/m^3. The concentration of NO_2 varies from 9.6 to 10.40 µg/m^3, which is within the prescribed limit of 80 µg/m^3. $PM_{2.5}$ shows variation in the concentration from 21.00 to 37.00 µg/m^3 which is within the permissible limit. The concentration of PM_{10} varies from 41.00 to 58.00 µg/m^3 which comes under the permissible limit.

TABLE 2. Concentration levels of $PM_{2.5}$, PM_{10}, SO_2 and NO_2 in the Khondbond iron ore mine and NAAQS

Location	Near Plant (A-1)	Mining Site (A-2)	Near Weighbridge (A-3)	Equipment Maintenance (A-4)	NAAQS
$PM_{2.5}$	27.6	26.5	23.25	14.8	60
PM_{10}	50.7	47.4	43.63	35.4	100
SO_2	9.63	9.59	9.61	9.34	80
NO_2	9.96	9.8	9.84	9.58	80

3.2. Mining site (A–2)

The second sampling point A–2 was not fixed it is a mobile point for sampling because excavation of ores will take place at the different location. The source of pollution in mining area is basically the movement of the vehicle, excavation of ores, drilling and blasting. The concentration of SO_2 varies from 9.20 to 10.00µg/m^3, which is within the permissible limit of 80µg/m^3. The concentration of NO_2 varies from 9.40 to 10.30µg/m^3, which is within the prescribed limit of 80 µg/m^3. $PM_{2.5}$ shows variation in the concentration from 16.00 to 33.00 µg/m^3 which is within the permissible limit. The concentration of PM_{10} varies from 37.00 to 53.00 µg/m^3 which comes under the permissible limit.

3.3. Weigh Bridge (A–3)

The main pollution near Weigh Bridge is vehicular pollution. Sampling point was at the side of Weigh Bridge. The RDS and FPM have been started at 6.0 AM and stopped at 6.0 AM next day. The concentration of SO_2 varies from 9.20 to 10.20 µg/m^3, which is within the permissible limit of 80 µg/m^3. The concentration of NO_2 varies from 9.40 to 10.30 µg/m^3, which is within the prescribed limit of 80 µg/m^3. $PM_{2.5}$ shows variation in the concentration from 16.00 to 30.00 µg/m^3 which is within the permissible limit. The concentration of PM_{10} varies from 35.00 to 50.00 µg/m^3 which comes under the permissible limit.

3.4. Equipment Maintenance (A–4)

The sampling point A-4 was nearer to the restroom the main source of pollution at equipment maintenance is vehicular pollution and population. The concentration of SO_2 varies from 9.00 to 9.70 µg/m^3, which is within the permissible limit of 80 µg/m^3. The concentration of NO_2 varies from 9.3 to 9.90 µg/m^3, which is within the prescribed limit of 80 µg/m^3. $PM_{2.5}$ shows variation in the concentration from 9.00 to 22.00 µg/m^3 which is within the permissible limit. The concentration of PM_{10} varies from 29.00 to 42.00 µg/m^3 which comes under the permissible limit.

The high concentration of PM_{10} at Near Plant (A–1) and Mining Site (A–2) was mainly due to different mining activities as well as running of vehicles on the unpaved road including abrasion of road materials, tires and brake linings as well as re-suspension of soil material because of traffic-induced turbulence [12] [13]. The sources of $PM_{2.5}$ have generally been confined to movement of the vehicle on paved/unpaved roads, vehicular exhaust (diesel-based), mining activities particularly drilling, blasting and crushing of rocks. The emission inventory indicates that heavy-duty diesel trucks were accountable for the majority of the exhaust particulate matter [14].

The main sources of NO_2 in mining rejoin are vehicular exhaust, blasting operations, etc. During combustion process (at high temperature) atmospheric nitrogen combines with oxygen to form NO_2 which is aggravated when the engine is diesel operated. Tunnel studies indicated that diesel engine produces five times the amount of NO_2 per mass of fuel burned when compared to gasoline vehicles [15].

3.5. Results of AQI

Data obtained from monitoring of ambient air at four different sites are used to calculate the sub-indices for critical parameters. The calculated AQI values for 24 hourly averages SO_2 and NO_2 concentrations are categorized as good during the study period at all the four sites. The AQI values calculated for $PM_{2.5}$ showed at equipment maintenance is under good category but for Near Plant (A-1), Mining site (A-2) and Near Weighbridge (A-2) under satisfactory category in case of PM_{10} the calculated AQI values lies under good category for A-2, A-3, and A-4 but lies under satisfactory category for A-1. The overall air quality was assessed using AQIs. The AQIs were calculated for criteria pollutants SO_2, NO_2, PM_{10} and $PM_{2.5}$ using IND-AQI procedure. The overall AQI was found to fall under the category of good.

TABLE 3. AQI and Sub-indices at different sampling sites

Location	Near Plant (A-1)	Mining Site (A-2)	Near Weighbridge (A-3)	Equipment Maintenance (A-4)
$PM_{2.5}$	46	44	39	25
PM_{10}	51	47	44	35
SO_2	12	12	12	12
NO_2	12	12	12	12
AQI	51	47	44	35

3.6. Air pollution control equipment and prevention techniques

As mechanized open cast iron ore mines are becoming larger, deeper and more capital intensive, resulting in environmental pollution. Continuing efforts have been made to improve ambient air quality condition upon the open cast mining activities. A management strategy was formulated for effective control of air pollution at source and other mitigative measures including green belt design have also been devised for sensitive areas [16]. Some pollution prevention measures and techniques were employed which are mobile water sprinkler, fixed water sprinkler, dry fog system at the primary crusher, dust extractor at the primary crusher, wet drilling, the addition of dust suppresscent in sprinkling water, wet processing of iron ore and transportation by conveyors.

4. Conclusion

The mineral industry in India is an important contributor to the country's GDP and foreign trade, and it also absorbs a considerable amount of the country's working population. To assess the environmental impact of different mining activities, SO_2, NO_2, PM_{10} and $PM_{2.5}$ were measured at all the four sites namely near the plant (A–1), at the mining site (A–2), near weighbridge (A–3), and at the site of Equipment Maintenance (A–4).

The analysis result illustrates particulate materials as the major pollutants to be concerned in mining areas. The ambient air concentration of PM_{10} concentrations at the monitoring station was well below the maximum allowed limit of National Ambient Air Quality Standards (NAAQS) for different areas. The high concentration of PM_{10} at Near Plant (A–1) and Mining Site (A–2) was mainly due to different mining activities as well as running of vehicles. According to the present study, NO_X and SO_2 concentrations at the monitoring station were well below the maximum allowed limit of National Ambient Air Quality Standards (NAAQS) for different areas. The AQIs were calculated for criteria pollutants SO_2, NO_2, PM_{10} and $PM_{2.5}$ using IND-AQI procedure. The overall AQI was found to fall under the category of good. Continuing efforts has been made to improve ambient air quality condition upon the open cast mining activities and many management strategies were formulated for effective control of air pollution at source.

References

[1] P. V. Mukunda Rao, V. Hima Bindu, G. Sagareshwar, J. Indracanti and Y. Anjaneyulu, Assessment of ambient air quality in the rapidly industrially growing Hyderabad urban environment, 2003.

[2] A. Masitah, H. Zaini and K. L. See, PM_{10} and Total suspended particulates (TSP) measurements in various power stations, *The Malayasian Journal of Analytical Sciences, 11(1),* 2007, 255-261.

[3] G. Singh and P. K. Sharma, Ambient air quality status in certain coal mining areas of Raniganj coalfield, *Energy Environment Monitor., 7(2),* 1991, 56-65.

[4] P. K. Sharma and G. Singh, Distribution of suspended particulate matter with trace element composition and apportionment of possible sources in the Raniganj Coalfields, India, *Env. Monit. Asses., 22,* 1992, 237-244.

[5] TERI, Area wide environmental quality management (AEQM) plan for the mining belt of Goa, Directorate of Planning statistics and Evaluation - Government of Goa, Goa, 1997.

[6] G. Singh, An index to measure depreciation in air quality in some coal mining areas of Korba industrial belt of Chhattisgarh, India, *Env. Monit. Asses., 122,* 2006, 309-317.

[7] D. Shooter and P. Brimblecombe, Air quality indexing, *International Journal of Environment Pollution, 36,* 2009, 305-323.

[8] F. Bruno and D. Cocchi, Recovering information from synthetic air quality indices, *Environmental Geology, 18,* 2007, 345–359.

[9] M. Sharma, Interpretation of air quality data using an air quality index for the city of Kanpur, India, *Journal of Environmental Engineering and Science, 2,* 2003, 453–462.

[10] M. Prakash and J. K. Bassin, Analysis Of Ambient Air Quality Using Air Quality Index – A Case Study, *International Journal of Advanced Engineering Technology, 1,* 2010, 106-114.

[11] NAAQS, National Ambient Air Quality Standard, India, prescribed by Central Pollution Control Board on 18 November, 2009.

[12] I. Barmapadimos, M. Nufer , D. C. Oderbolz, J. Keller, S. Aksoyoglu, C. Hueglin, U. Baltensperger and A. S. H. Prevot, The weekly cycle of ambient concentration and traffic emission of coarse (PM_{10} to $PM_{2.5}$) atmospheric particles. *Atmospheric Environment, 45,* 2011, 4580-4590.

[13] N. Bukowiecki, P. Lienemann, M. Hill, M. Furger, A. Richard, F. Amato, A. S. H. Prevot, U. Baltenspurger, B. Buchmann, and R. Gehrig, PM_{10} emission factors for non-exhaust particles generated by road traffic in an urban street canyon and along a freeway in Switzerland. *Atmospheric Environment, 44,* 2010, 2330-2340.

[14] R. F. Sawyer, R. A. Harley, S. H. Cadle, J. M. Norbeck, R. Slott and H. A. Bravo, Mobile sources critical review: 1998 NARSTO assessment. *Atmospheric Environment, 34,* 2010, 2161-2181.

[15] W. T. Kirchstetter, H. A. Miguel and A. R. Harley, On road comparison of exhaust emission from gasoline and diesel engines. In: Proceedings of the Eighth CRC On-road Vehicle Emission Workshop, 1998, April 20-22,San Diego, CA.

[16] R. Trivedi, M. K. Chakrabotry and B. K. Tewary, Dust dispersion modeling using fugitive dust model at an opencast coal project of Western Coalfields Limited, India, *Journal of scientific and Industrial research, 68,* 2009, 71-78.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325145142

Study of Improvement in the Performance of Triple Effect Vapor Absorption System Using Loop Heat Pipes

Article · January 2018

CITATIONS
0

2 authors, including:



Ankit Dwivedi
Delhi Technological University
8 PUBLICATIONS **0** CITATIONS

SEE PROFILE

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

Doctoral Research View project

All content following this page was uploaded by Ankit Dwivedi on 15 May 2018.

The user has requested enhancement of the downloaded file.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

Study of Improvement in the Performance of Triple Effect Vapor Absorption System Using Loop Heat Pipes

Ankit Dwivedi[1], R S Mishra2
[1,2]*Dept. of Mechanical Engineering, Delhi Technological University, Delhi*

Abstract: ***The triple effect vapor absorption refrigeration system (VARS) works at higher temperatures (>150-160°C).This high temperature working of the cycle helps in achieving higher COP $_I$ and COP $_{II}$. This research analysis works on reducing the heat input in the generators by use of two Loop Heat Pipes (LHPs) installed in the high and medium temperature levels. The simulations have found that the improved COP $_I$ &COP $_{II}$ are 2.545 and 0.588 respectively. Also the percentage increase in performances of the modified system has been recorded as 61.112 % and 25.146% respectively. Simulations also found that the average T_C can be maintained at 151.18 °C. The increase in the performance characteristics shows a positive gradient with the increase in the temperature of operation of LHP. The average Q_{Leak} and Q_{Cond} are found to be 89.44 kW and 71.68 kW respectively***.

I. INTRODUCTION

Loop Heat Pipe (LHP) is an evaporator-condenser that works on the thermo-siphon effect. It operates and transfers heat without any requirement of input work. It acts as a super-conductor owing to the high heat transfer rates related to the evaporation and condensation. The heat pipe is a wicked structure, outside of which the heat is absorbed by the working fluid of the LHP from the ambient and the fluid gets vaporized. This vapor passes through the porous wicked structure and moved towards the condenser of the LHP through the Vapor Line. The vapor rejects heat in the condenser and gets condensed. These condensates then move towards the Compensation Chamber (CC) connected to the evaporator of the LHP and gets accumulated there.

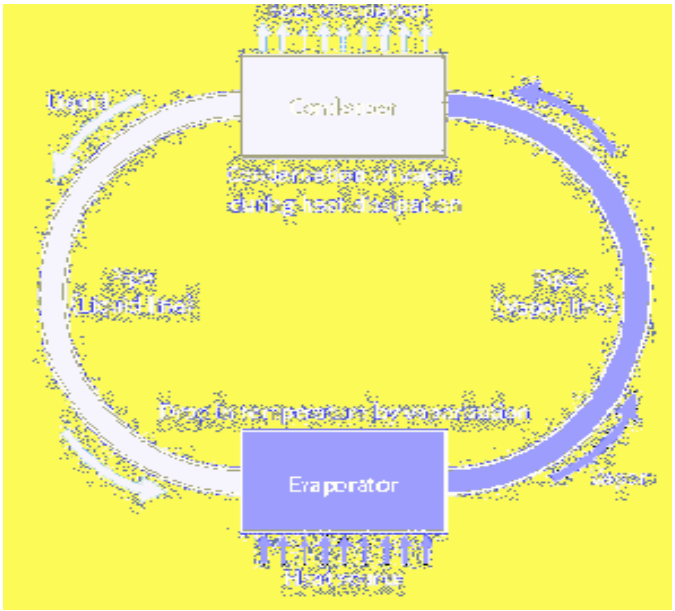


Fig 1: Cyclic process of a Loop Heat Pipe [35]

There is some leakage of heat of associated with the LHP from CC, Liquid and Vapour Lines, From the wick etc. These losses are although negligible but can be reduced up to an extent by the proper selection of materials and designs.

©IJRASET (UGC Approved Journal): All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

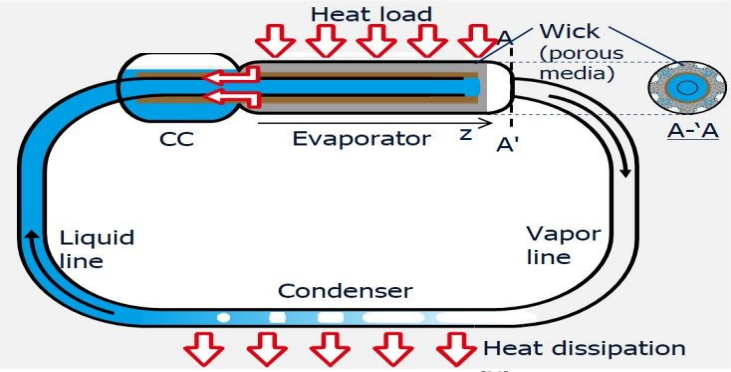


Fig 2: Schematic of LHP [36].

A triple effect vapor absorption system refrigeration system (VARS) is used when the source temperature is high(>140-150 °C). The high temperature also corrodes the material of the generator and connecting lines. The parallel flow type heat exchanger is shown and explained in fig 3. This VARS system is very useful in utilizing a waste heat. The efficiency of the triple effect system is also relatively higher as the operating temperature is very high comparatively.

II. LITERATURE REVIEW

Fabian Korn et al. [2012]performed several vital experiments on heat pipes to establish it to be one of the most effective procedures to transport thermal energy from one point to another, mostly used for cooling[6].Sameer Khandekaret al.[2010]performed experiments on the global thermal performance modeling of Pulsating Heat Pipes (PHPs) requires local, spati-o-temporally coupled, flow and heat transfer information during the characteristic, self-sustained thermally driven oscillating Taylor bubble flow, under different operating conditions[7].JozefHužvár, PatrikNemecet al. [2007]used heat pipe, observed its basic principles and operating limits. High temperature heat pipes were evaluated for use in energy conversion applications such as fuel cells, gas turbine re-combustors, and Stirling cycle heat sources, with the resurgence of space nuclear power, additional applications include reactor heat removal elements and radiator elements[8].R.Z. Wanget al.[2008]added heat pipes in adsorption water chiller or ice maker initials. His work showed that the adsorption refrigerators are very efficient [10].PrachaYeunyongkul et al.[2009]aimedat experimentally investigating the application of a closed loop oscillating heat pipe (CLOHP) as the condenser for a vapor compression refrigeration system[14].R. Rajashree et al.[1990]went through a numerical analysis of an unsteady, viscous, laminar, incompressible, two dimensional heat and mass transfer, in the vapour gas region of gas loaded circular heat pipe [20]. Da-Wen Sun (1996) performed a detailed thermodynamic analysis of the properties of these binary fluids and expressed in polynomial equations. The performances of three cycles were compared. M.M. Talbi et al. (2000) carried out an exergy analysis on a single-effect absorption refrigeration cycle with lithium-bromide±water as the working Fuid pair. E. Kurem et al.(2001) analyzed the Absorption Heat Pump (AHP) and Absorption Heat Transformers (AHT) using ammonia-water and water-lithium bromide solutions. A fundamental AHP and AHT systems was described and explained the operating sequence. R.D. Misra et al. (2002) applied the therm-o-economic theory is to the economic optimization of a single effect water/LiBr vapour absorption refrigeration system for air-conditioning application. S.A. Adewusi et al (2004). studied the performance of single-stage and two-stage ammonia–water absorption refrigeration systems (ARSs).They calculated entropy generation of each component and the total entropy generation of all the system components as well as COP of the ARSs . S. Arivazhagan et al. (2006) investigated experimentally on the performance of a two-stage half effect vapour absorption cooling system .The prototype is designed for 1 kW cooling capacity using HFC based working fluids (R134a as refrigerant and DMAC as absorbent). RabahGomri et al. (2008) performed exergy analysis of double effect lithium bromide/water absorption refrigeration system. The system consisted of a second effect generator between the generator and condenser of the single effect absorption refrigeration system, including two solution heat exchangers between the absorber and the two generators. S.C. Kaushik et al. (2009) presented the energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption systems. They developed a computational model for the parametric investigation of the systems. Berhane H. Gebreslassie et al. (2010) performed an exergy analysis, which only considered the unavoidable exergy destruction, conducted for single, double, triple and half effect Water–Lithium bromide absorption cycles. Gulshan Sachdeva et al.(2014) performed anexergy analysis of VAR system using LiBr-H2O as working fluid with the modified Gouy-Stodola approach. Karl Ochsner et al. (2008) developeda new CO2-heat pipe with high-grade steel corrugated pipe system, which – contrary to other

©IJRASET (UGC Approved Journal): All Rights are Reserved

pipe systems permits raw length up to 100 m. They also described the establishment of the heat pump system in general. Ankit Dwivedi et al. (2018) through computer simulations of the replacement of condensers from single, double and half effect VARS, showed the improvements in the performance in COP $_{I}$ and COP $_{II}$ .Waste heat is being utilized in the cycle from some outside source, while the cycle itself rejects heat at the saturation temperature of the refrigerant which in case is as high as the generator. This heat is taken away by the working fluid of the condenser (Air, Water etc). It can be taken away by the mixture of refrigerant from the absorber as well which already works at the condenser temperature. The LHP can be used for this purpose. It can replace the condenser and help in the intra cycle heat exchange. The two condensers working at higher temperatures (>100°C) can be replaced in the triple effect cycle by LHPs. The bulky condensers will be removed and flexile heat pipes will take its place. Also as mentioned earlier the LHPs will act as a superconductor with having higher capacity and higher heat transfer coefficient. Some of the research works have showed the options for more analysis in these aspects.

III. SYSTEMS DESCRIPTION

It can be explained by the fig 3 what are the basic components and cycle processes of the triple effect VARS. This is a parallel flow triple effect system which has 3 condensers (Low(LC), Medium(MC) and High(HC)), 3 generators (Low(LG), Medium(MG) and High(HG)), an absorber, an evaporator, 3 Heat Exchangers (Low H.E x, MedH. Ex, HighH. Ex), pumping devices and throttle valves. The mixture is pumped to the high temperature generator HG (>working at 150°C) through the 3 heat exchangers. A major part of the rich refrigerant then gets separated from the mixture and is transferred through HC to the evaporator to produce the refrigeration. The remaining mixture from the HG then is throttled to the MG(>120°C) where it is heated and the separated strong refrigerant is transferred to the evaporator through MC and throttle valves. The remaining mixture from MG is then throttled to the LG(>80-90°C) where the above process takes place. Finally the remaining mixture is throttled to the absorber. This 3 stage process helps the separation strong refrigerant well administered.

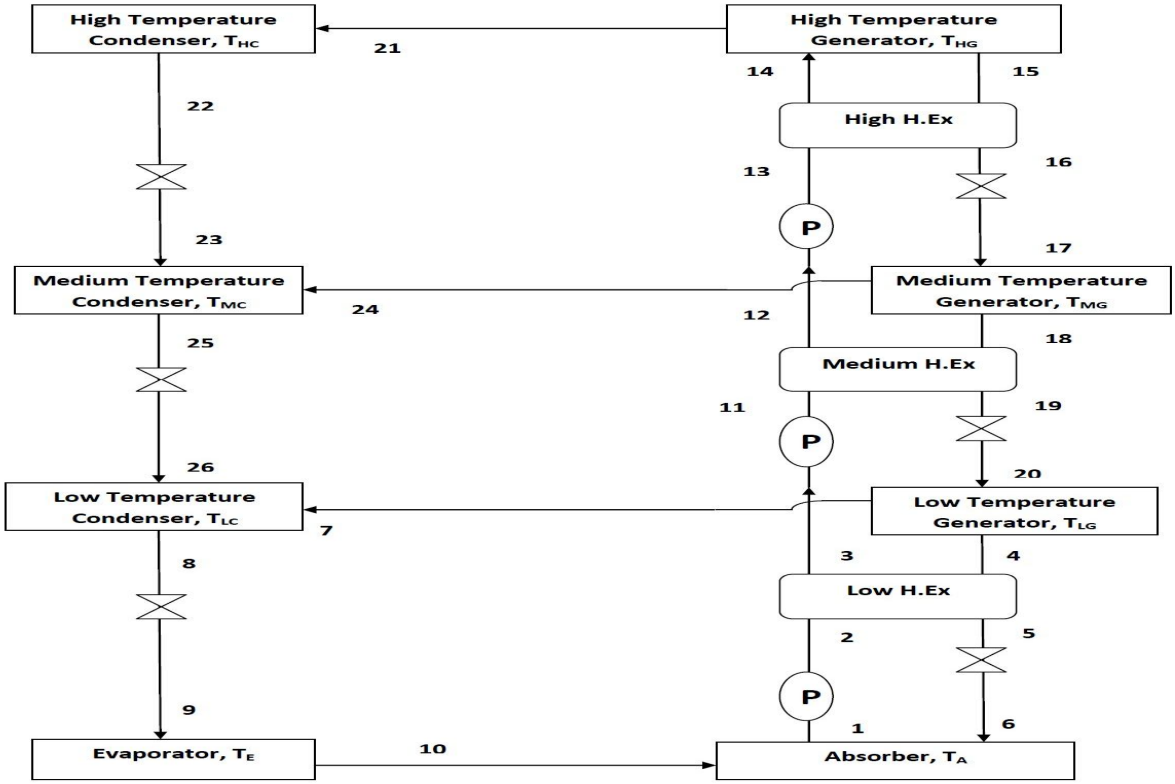


Fig 3: A Triple Effect (Parallel) Vapour Absorption System

The proposed modifications can be well described in fig 4. Here HC, LC , Low H.Ex and Med H.Ex are to be removed from LHPs. Understanding that the temperature is sufficiently high at this level(i.e. 120°C -150°C), LHPs will perform the tasks of condensers and H.Ex reduces the size and complexity of the Heat Exchangers. The cost associated can also be reduced. The design of the LHP

©IJRASET (UGC Approved Journal): All Rights are Reserved

will be very simple when compared with that of H. Ex. Heat is removed at the evaporator size on the LHP from the condensing strong refrigerant. This heat will be transferred to the mixture coming from the absorber. The effectiveness on the the LHP will be higher than the other heat exchangers owing to the high heat transfer coefficients. The simulations are executed and the results are as follows. The table 1 consists of the terminology used in the analysis.

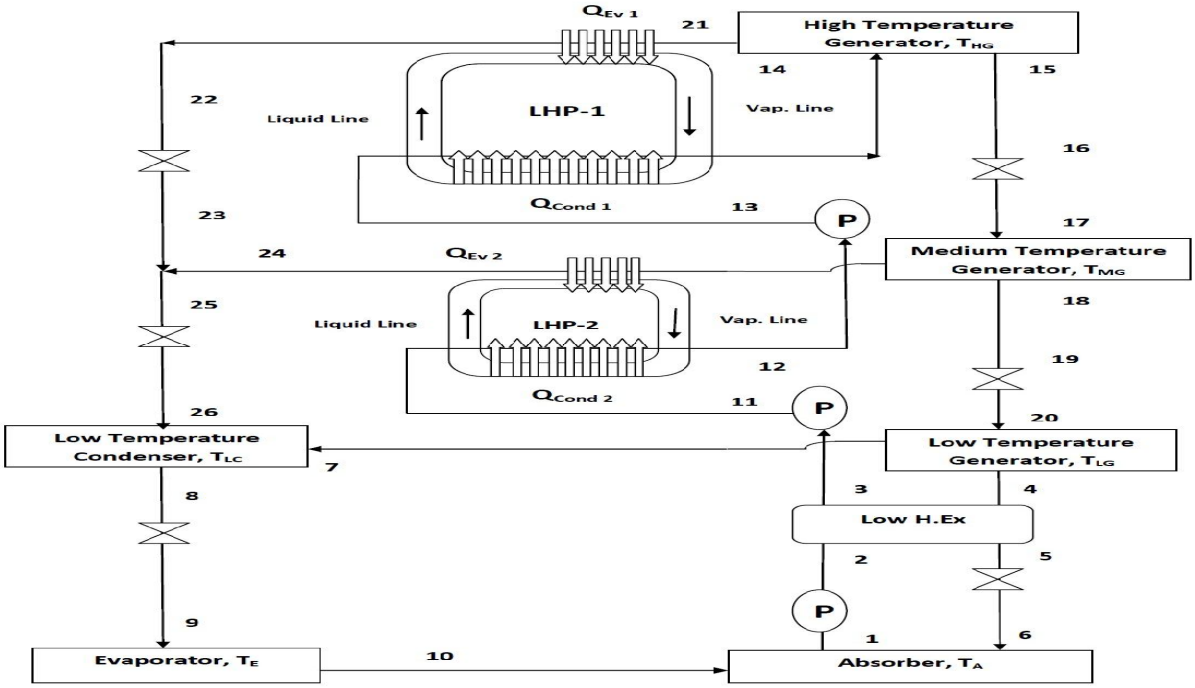


Fig.4: Modified Triple Effect Vapour Absorption System

Table 1: Terms Used in Simulation

Terms	Abbreviations
Refrigeration Effect in kW	RE (kW)
Heat rejected in absorber in kW	Q_a(kW)
Heat supplied in generator in kW	Q_g (kW)
Heat rejected in condenser of LHP in kW	Q_{cond}(kW)
Heat absorbed in evaporator of LHP in kW	Q_{eva}(kW)
Absorber Temperature in°C	T_{La},T_{Ha}(°C)
Generator Temperature in °C	T_{Hg}, T_{LG},T_{MG}(°C)
LHP Condenser Temperature in °C	T_c (°C)
Evaporator Temperature in °C	T_E,T_e(°C)
Heat Rejected in Condenser in kW	Q_C (kW)
First Law Coefficient of Performance	COP_I
Second Law Coefficient of Performance	COP_{II}
Heat Leaked from the LHP in kW	Q_{Leak} (kW)
Percentage Improvement in First Law Coefficient of Performance	$\%COP_{I\ imp}$
Percentage Improvement in Second Law Coefficient of Performance	$\%COP_{II\ imp}$
Improvement in First Law Coefficient of	$COP_{I\ imp}$

©IJRASET (UGC Approved Journal): All Rights are Reserved

Performance	
Improvement in Second Law Coefficient of Performance	$COP_{II\ imp}$
Low Temperature Generator, Condenser	LG,LC
Medium Temperature Generator, Condenser	MG, MC
High Temperature Generator, Condenser	HG, HC

IV. RESULTS AND DICSUSSIONS

From Fig 5 to Fig 8 shows the various variations in COP_I & COP_{II}. Fig 5 describes the comparison of COP_I improvement in the modified system and the COP_I in the basic triple effect cycle. The average COP_I of modified system can be observed to be 2.545.

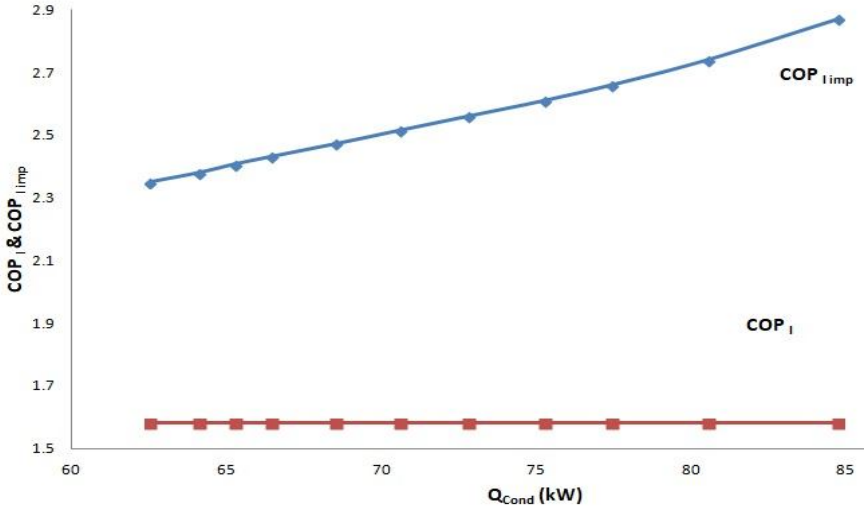


Fig 5: Comparison between COP_I and $COP_{I\ imp}$ plotted with Q_{Cond}

It can be seen in the fig 6 how COP_{II} of the proposed modified system stands against the original system. There's a continuous increase in the COP_{II}, as the slope in the beginning is positive and in the later stage the increase in Q_{Cond} may not yield in higher rises in COP_{II}. The average COP_{II} of modified system can be 0.588.

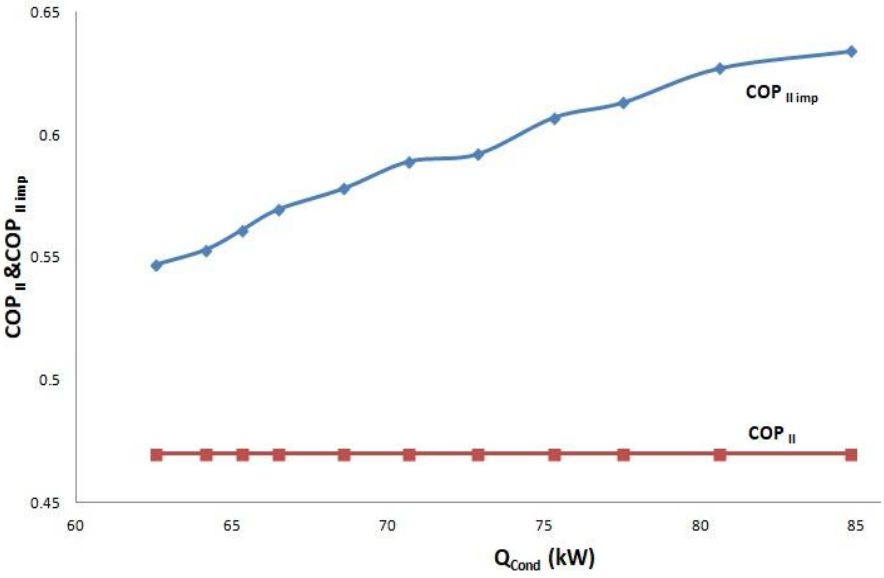


Fig 6: Comparison between COP_{II} and $COP_{II\ imp}$ plotted with Q_{Cond}

©IJRASET (UGC Approved Journal): All Rights are Reserved

Fig 7 shows the increase in COP $_{I}$& COP $_{II}$ relatively over the range of Q $_{Cond}$. It can be seen that COP $_{I}$ has a higher and direct effect of Q $_{Cond}$ as its being used in improving it directly. The COP $_{II}$ on the other hand doesn't show a sharp rise as COP $_{I}$ does. The reduction of heat wastage in the system no doubt helps in decreasing the exergy loss. Both external and internal irreversibility are reduced.

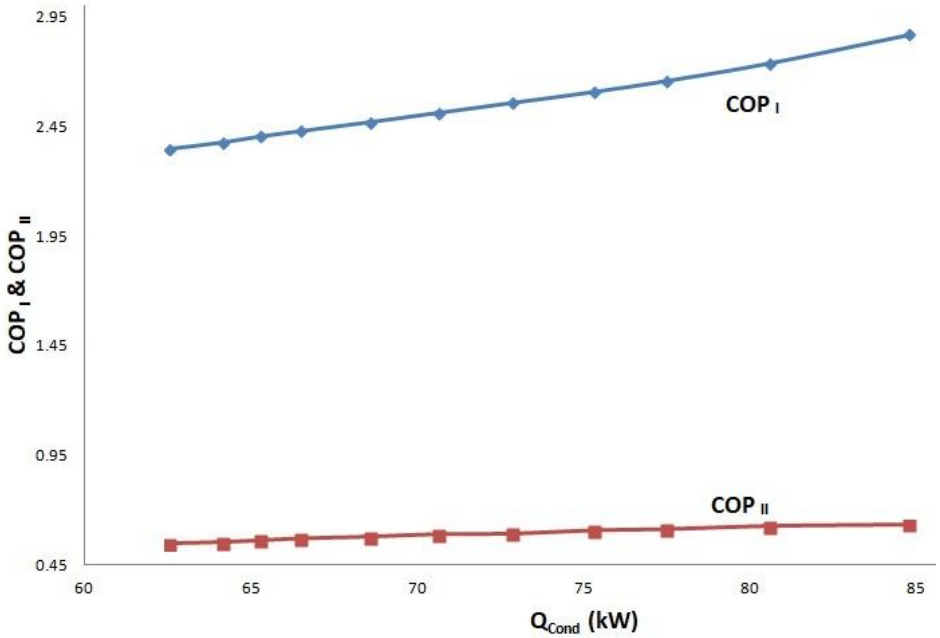


Fig 7: COP $_{I}$ and COP $_{II}$plotted with Q_{Cond}

Fig 8 in a way is developed to show the percentage rise in the COP_{I} and COP $_{II}$ owing to the use of LHPs. As discussed earlier the COP $_{I}$ is directly affected by the Q $_{Cond}$. Here the average percentage rise in COP $_{I}$is found to be around 61.112 % and the average rise in COP_{II} is 25.146%. Also if the designs can be changed so that LHPs are able utilize more heat and the leaks are reduced the COP $_{I}$ can be further improved.

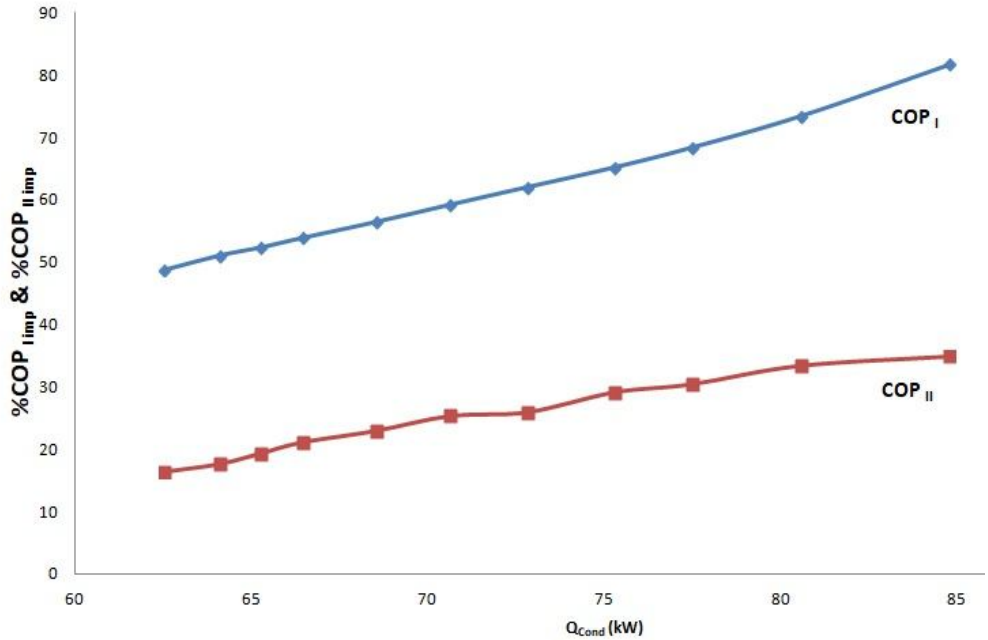


Fig8: Variation of %COP $_{I\,imp}$ and % COP $_{II\,imp}$ plotted with Q_{Cond}

©IJRASET (UGC Approved Journal): All Rights are Reserved

The Fig 9 & Fig 10 shows the percentage improvements in the performance with the range of LHP condenser temperature T_C and HG temperature T_G. With the ranging T_C rise in COP_I has a positive slope as discussed earlier and as the T_C increase further the availability also increases. Increase in the availability helps in increasing COP_I as the heat to be utilized increases. COP_{II} also follows, and increase in the COP_{II} is recorded by the simulations.

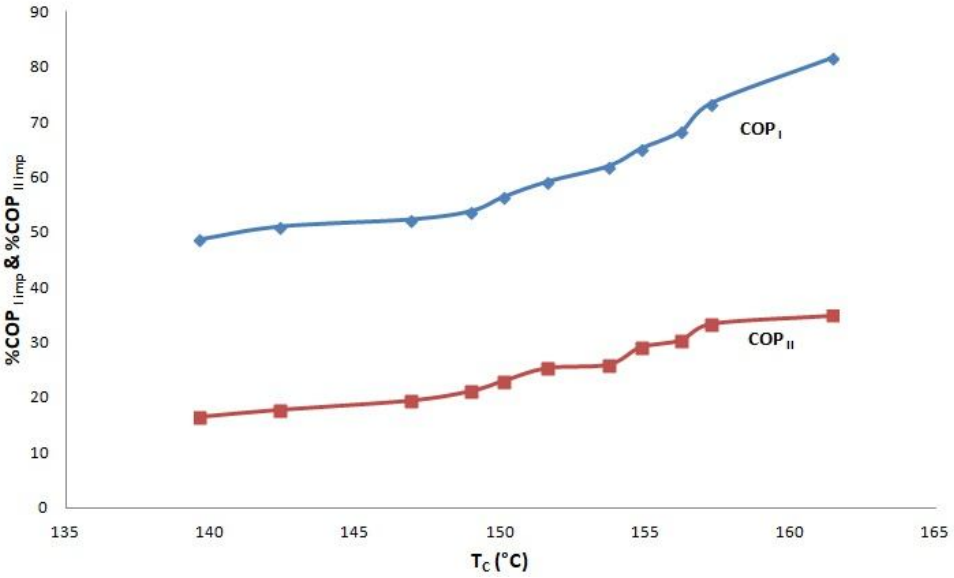


Fig9: Comparison of %age improvements in COP_I & COP_{II} varying the T_C

Fig 10 describes the variations of performance with the range of T_G. With rise in T_G the COP_I has a sharp increase, whereas the COP_{II} has a gradual increase.

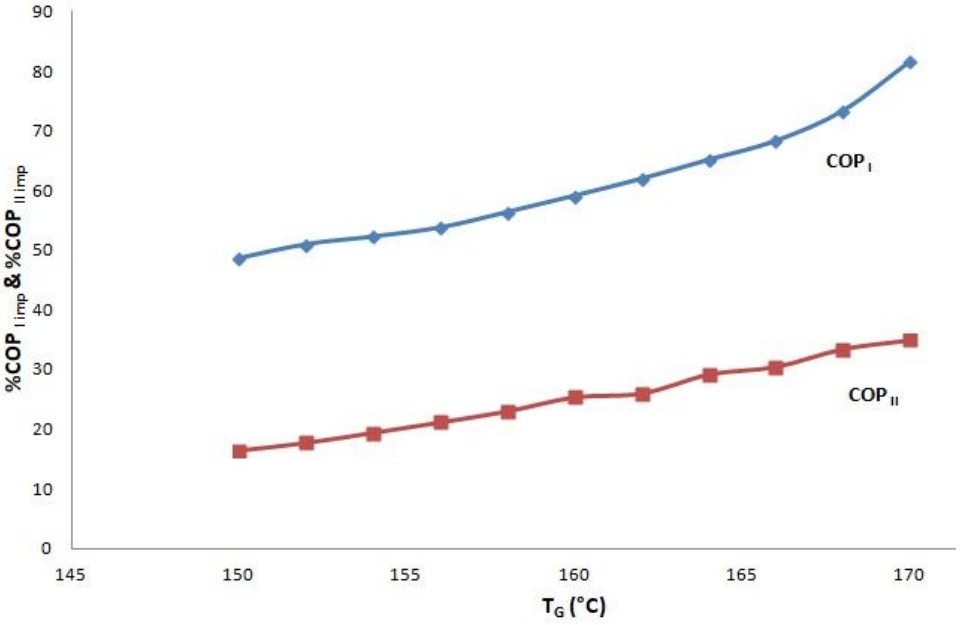


Fig 10: Comparison of %age improvements in COP_I & COP_{II} varying the T_G

The Fig 11 & Fig 12 show the improved performance with the temperature range.

©IJRASET (UGC Approved Journal): All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

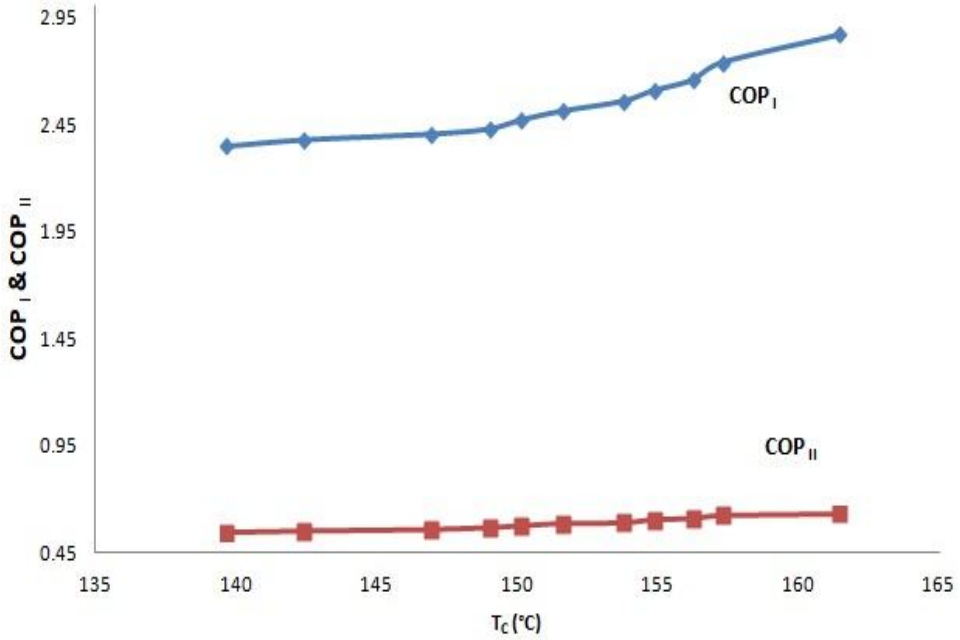


Fig 11: Comparison of COP $_{I}$& COP $_{II}$ varying the T_C

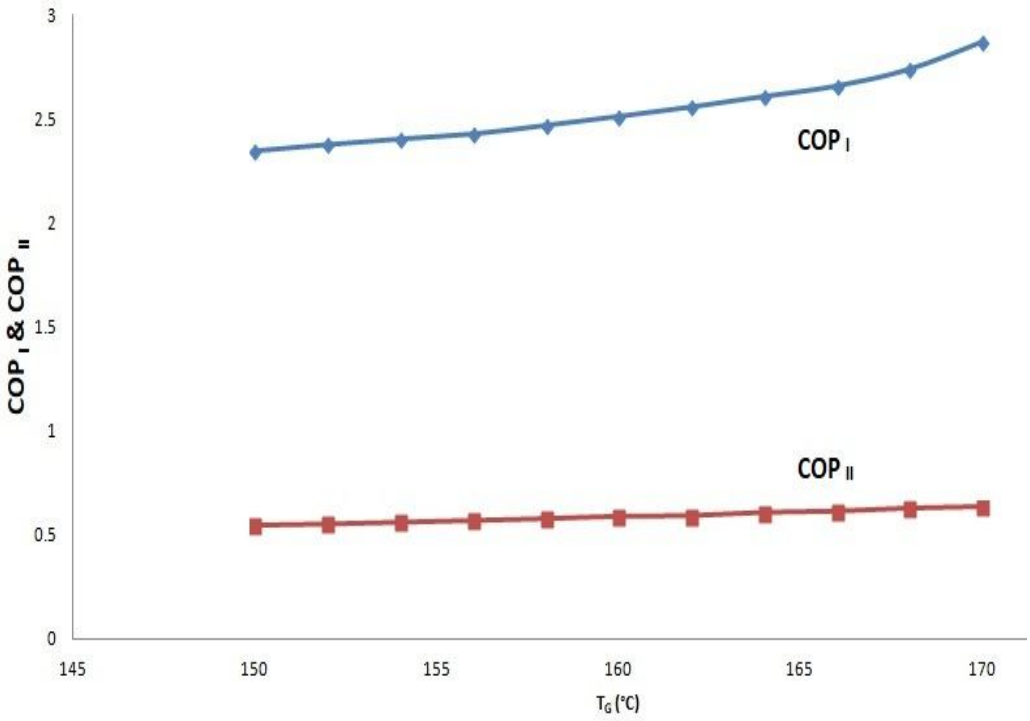


Fig 12: Comparison of COP $_{I}$& COP $_{II}$ varying the T_G

Fig 13 and Fig 14 describe the variations in Heat Leaked and Heat Utilized in the LHPs along the temperature. In the figure as the temperature of the LHP condenser increases it can be observed that heat utilized in condenser increases. The average heat utilized in the LHP condenser is 71.68 kW.

©IJRASET (UGC Approved Journal): All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

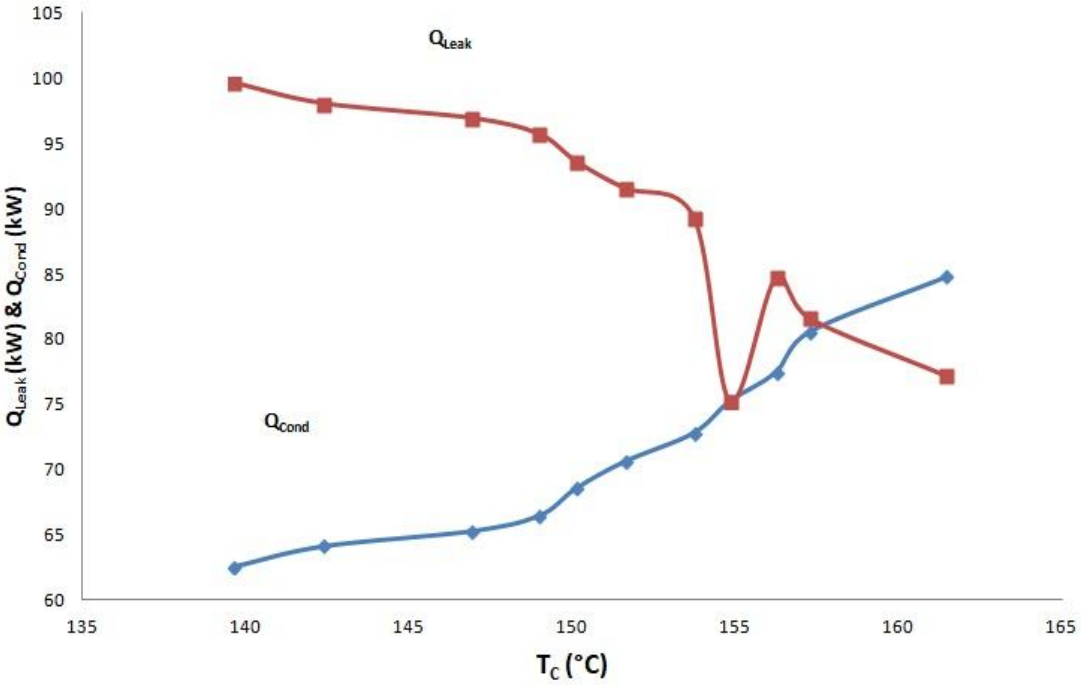


Fig 13: Comparison of Q_{Leak} & Q_{Cond} varying the T_C

On the other hand, the figure below shows the same variation with temperature of HG. The heat leak Q_{Leak} is decreasing with the increase in T_G. The average heat that has leaked the LHP is 89.44 kW over the ranges of temperatures.

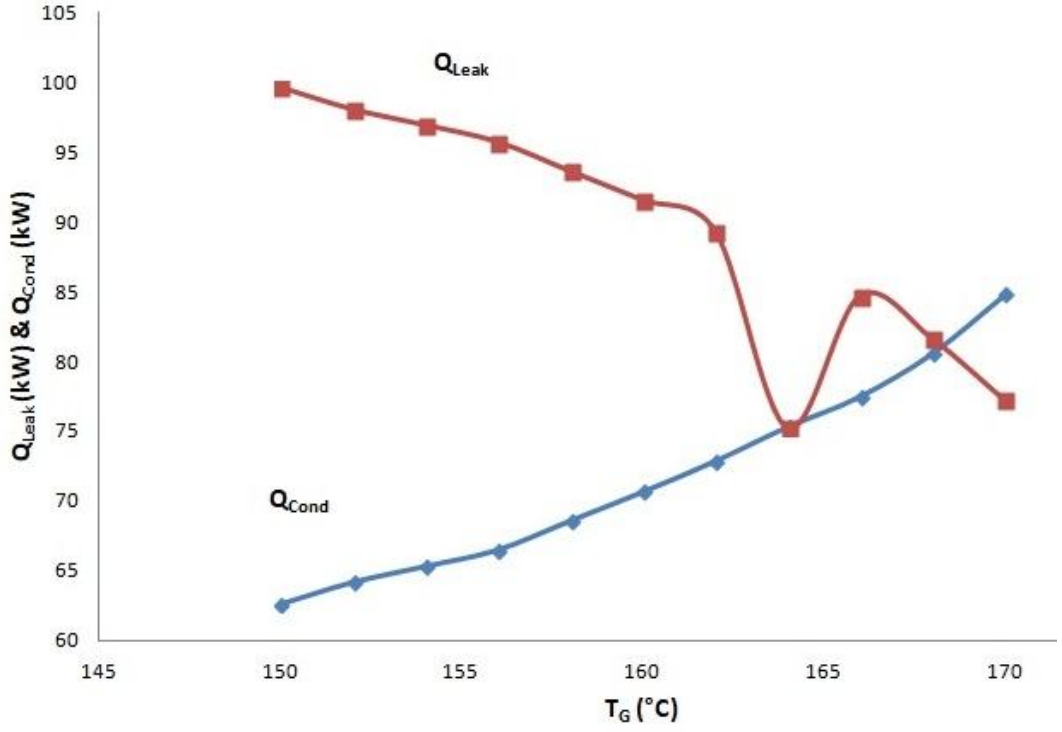


Fig 14: Comparison of Q_{Leak}&Q_{Cond} varying the T_G

Fig 15 shows the dependency of TC on TG.As the TG increases TC also increases. Direct dependence can be observed from the figure below. The average temperature can be managed to be achieved in the condenser of LHP is 151.18 °C.

©IJRASET (UGC Approved Journal): All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

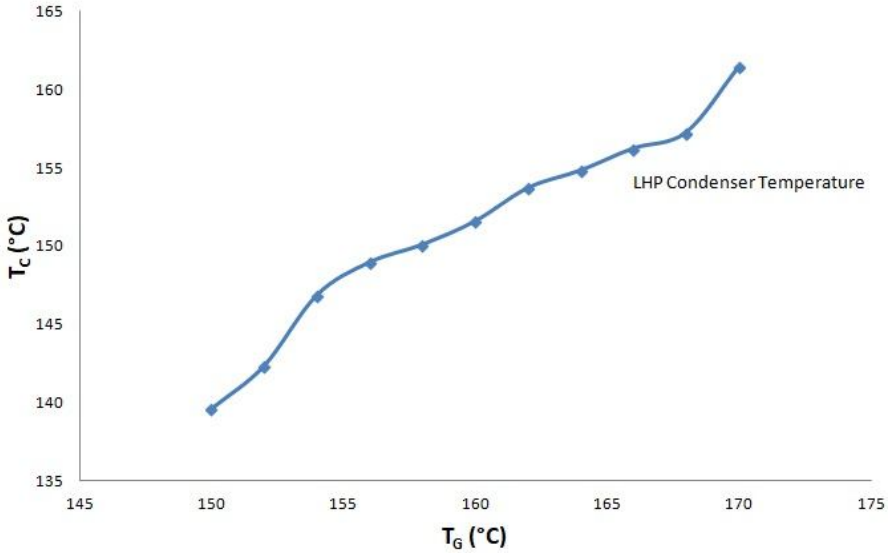


Fig 15: Variation of T_C with the T_G

The Figure 16 to 19 the performance of the modified system can be studied with the variations in the working temperatures.

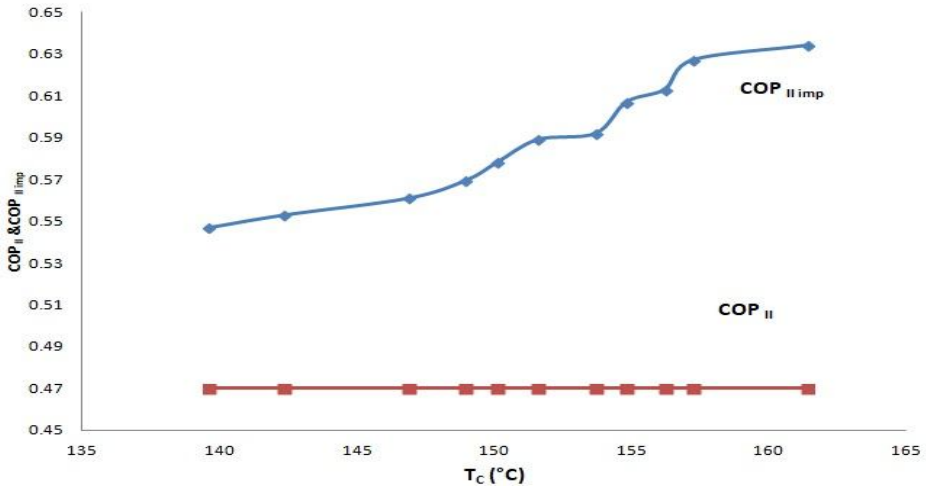


Fig 16: Variation of improvement in COP_{II} with the T_C

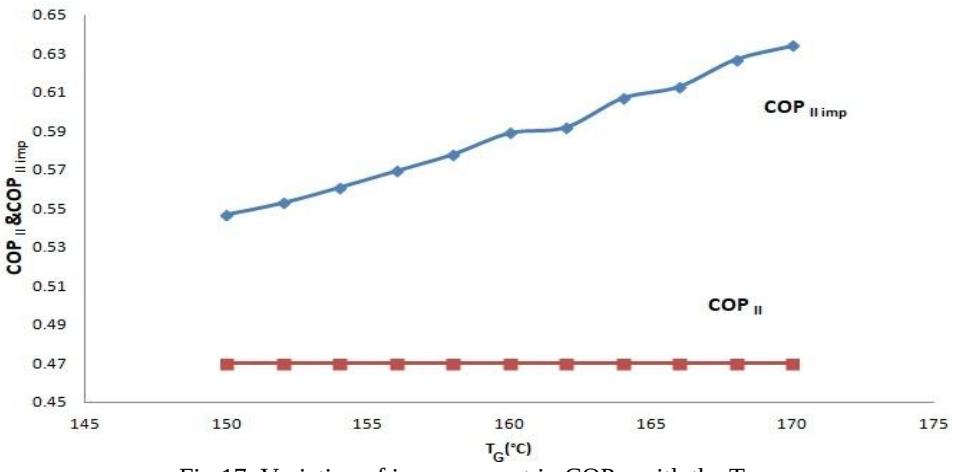


Fig 17: Variation of improvement in COP_{II} with the T_G

©IJRASET (UGC Approved Journal): All Rights are Reserved

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

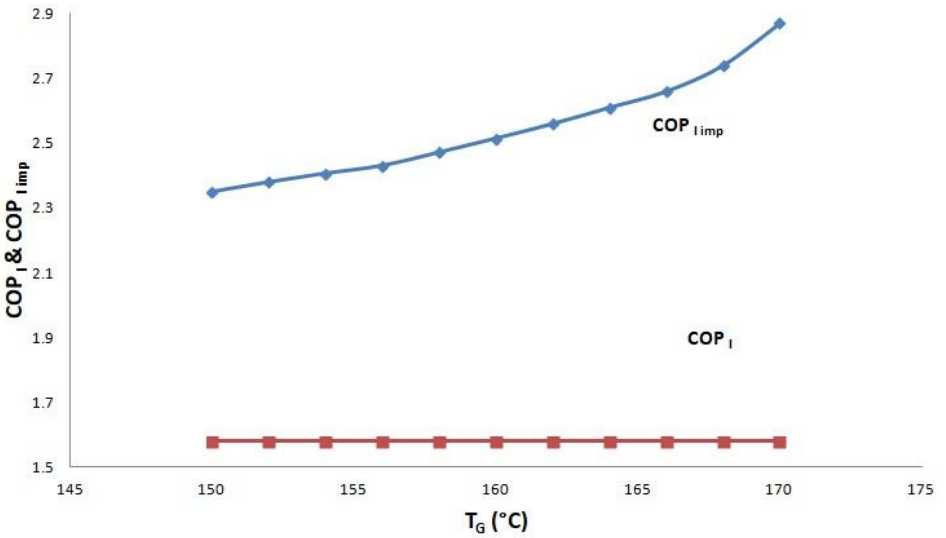


Fig 18: Variation of improvement in COP $_I$ with the T_G

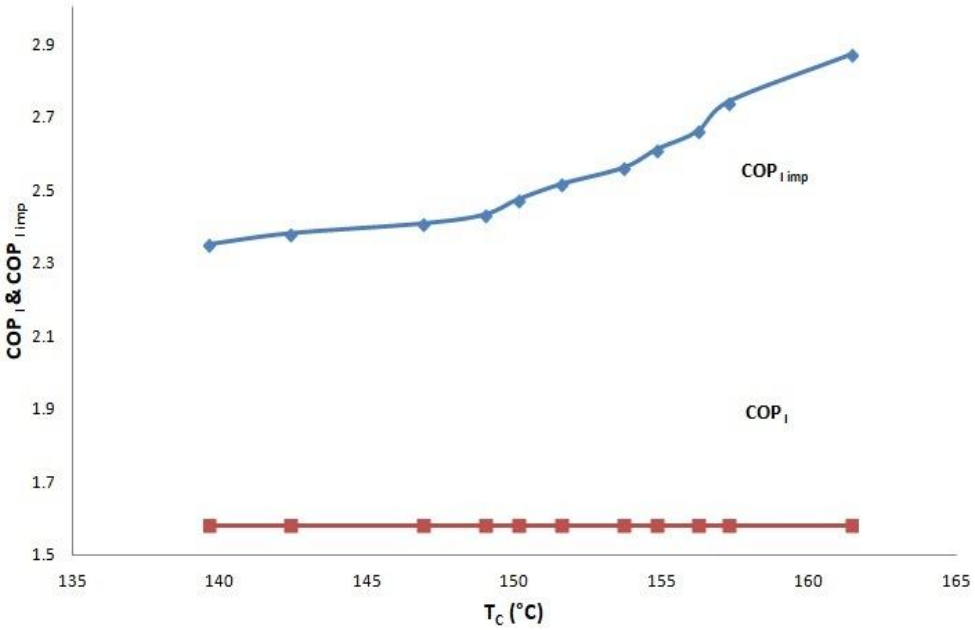


Fig 19: Variation of improvement in COP $_I$ with the T_C

V. CONCLUSIONS

Following the results of the simulations, following conclusions can be made

A. The average improved COP $_I$ &COP $_{II}$ are 2.545 and 0.588 respectively.

B. The percentage increase in COP $_I$ &COP $_{II}$ can be observed to be 61.112 % and 25.146% respectively.

C. The COP $_I$ has a sharper rise over the range when compared to the COP $_{II}$ for the range of Q_{Cond}.

D. The COP $_I$ shows an increasing slope for the entire range while the COP $_{II}$ shows a decreasing slope for the entire range of the T_G and T_C.

E. Increase in the temperature T_G provides the desirable results such as reduction in the Q_{Leak} and increase in Q_{Cond} and average T_C is found to be is 151.18 °C.

F. Average Q_{Leak} and Q_{Cond} are found to be 89.44 kW and 71.68 kW respectively.

©IJRASET (UGC Approved Journal): All Rights are Reserved

REFERENCES

[1] Saeed. Sedigh , Hamid. Saffari ,"Thermodynamic analysis of single effect and half effect absorption refrigeration system" International Journal of Energy & Technology Vol.25 (2011) 1-9
[2] S. Arivazhagan , R. Saravanan , S. Renganarayanan ," Experimental studies on HFC based two-stage half effect vapour absorption cooling system" Applied Thermal Engineering Vol. 26 (2006) 1455–1462
[3] ulshanSachdeva, Ram Bilash," Thermodynamic Analysis of a Vapor Absorption System Using Modified Gouy-Stodola Equation" International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:8, No:12, 2014
[4] I. Horuz," A comparison between Ammonia-water and Water-Lithium Bromide solutions in Vapor Absorption Refrigeration Systems" Int. Comm. Heat Mass Transfer, Vol. 25, No. 5, pp. 711-721, 1998
[5] Abdul Khaliq, and Rajesh Kumar," Exergy analysis of double effect vapor absorption refrigeration system" Int. J. Energy Res. 2008; Vol.32:161–174
[6] S.C. Kaushika, AkhileshArora," Energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption refrigeration systems" international journal of refrigeration Vol.32 (2009) 1247 – 1258
[7] RabahGomri , RiadHakimi," Second law analysis of double effect vapour absorption cooler system" Energy Conversion and Management Vol.49 (2008) 3343–3348
[8] S.A. Adewusi, Syed M. Zubair," Second law based thermodynamic analysis of ammonia–water absorption systems" Energy Conversion and Management Vol. 45 (2004) 2355–2369.
[9] RabahGomri," Second law comparison of single effect and double effect vapour absorption refrigeration systems" Energy Conversion and Management Vol. 50 (2009) 1279–1287
[10] M. Belghazi, A. Bontemps, C. Marvillet," Experimental study and modelling of heat transfer during condensation of pure fluid and binary mixture on a bundle of horizontal finned tubes" International Journal of Refrigeration Vol.26 (2003) 214–223.
[11] M.M. Talbi, B. Agnew," Exergy analysis: an absorption refrigerator using lithium bromide and water as the working Fuids" Applied Thermal Engineering Vol. 20 (2000) 619-630
[12] E. Kurem," A comparison between Ammonia-water and Water-Lithium Bromide solutions in vapour absorption heat transformers "In. Comm. Heat Mass Transfer; Vol. 28, No. 3, pp. 421-438, 2001
[13] Da-Wen Sun," Comparison of the performances and NH3-H20, NH3-LiNO3 and NH3-NaSCN VaporAbsorption Refrigeration Systems" Energy Convers. Mgmt Vol. 39, No. 5/6, pp. 357-368, 1998
[14] Yu.F. Maydanik," Review Loop heat pipes" Applied Thermal Engineering Vol.25 (2005) 635–657
[15] Randeep Singh, AliakbarAkbarzadeh , MasatakaMochizuk," Operational characteristics of a miniature loop heat pipe with flat evaporator" International Journal of Thermal Sciences Vol.47 (2008) 1504–1515
[16] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, C.J. Chen," Performance study of a high efficient multifunction heat pipe type adsorption ice making system with novel mass and heat recovery processes" International Journal of Thermal Sciences Vol.46 (2007) 1267–127
[17] Yuan-Ching Chiang , Wen-Cheng Kuo , Chia-CheHo , Jen-JieChieh," Experimental study on thermal performances of heat pipes for air-conditioning systems influenced by magnetic nanofluids, external fields, and micro wicks" International Journal of Refrigeration Vol.43 (2014) 62 -70
[18] T.X. Li, R.Z. Wang, L.W. Wang, Z.S. Lu," Experimental investigation of an innovative dual-mode chemisorption refrigeration system based on multifunction heat pipes" International Journal of Refrigeration 3 1 (2 0 0 8) 1 1 0 4 – 1 1 1 2
[19] L. GarousiFarshi a,*, C.A. Infante Ferreira b, S.M.S. Mahmoudi a, M.A. Rosen," First and second law analysis of ammonia/salt absorption refrigeration systems" International Journal of Refrigeration Vol. 4 0 (2 0 1 4) 1 1 1-1 2 1
[20] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, J.Y. Wu," Influence of mass recovery on the performance of a heat pipe type ammonia sorption refrigeration system using CaCl2/activated carbon as compound adsorbent" Applied Thermal Engineering Vol. 28 (2008) 1638–1646
[21] Z.S. Lu , L.W. Wang, R.Z. Wang," Experimental analysis of an adsorption refrigerator with mass and heat-pipe heat recovery process" Energy Conversion and Management Vol. 53 (2012) 291–297
[22] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences Vol.50 (2011) 592-598
[23] Basant K. Agrawal*, Munawar N. Karimi," Thermodynamic performance assessment of a novel waste heat based triple effec refrigeration cycle" International Journal of Refrigeration Vol. 3 5 (2 0 1 2) 1 6 4 7-1 6 5 6
[24] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences 50 (2011) 592-598
[25] T.S. Jadhav , M.M. Lele," Theoretical energy saving analysis of air conditioning system using heat pipe heat exchanger for Indian climatic zones" Engineering Science and Technology, an International Journal Vol. 18 (2015) 669-673
[26] Chengchu Yan , Wenxing Shi , Xianting Li , Shengwei Wang," A seasonal cold storage system based on separate type heat pipe for sustainable building cooling" Renewable Energy Vol.85 (2016) 880-889
[27] Matthias H. Buschmann," Nanofluids in thermosyphons and heat pipes: Overview of recent experiments and modelling approaches" International Journal of Thermal Sciences Vol.72 (2013) 1-17
[28] P.D. Dunn, D.A. Reay, Heat Pipes, Pergamon Press, Oxford, 1993
[29] D.Reay, Heat Pipes-Theory, Design and Applications, Butterworth-Heinemann,Oxford, Fifth edition 2006
[30] Korn, F., "Heat Pipes and its Applications" Project Report 2008 MVK160 Heat and Mass Transport May 07, 2008, Lund, Sweden
[31] Rajashree, R., Rao, K.S., "**A** Numerical Study of the Performance of Heat Pipe" Indian Journal of Pure and Applied Mathematics, 21 (1): 95-108, January 1990
[32] C.P. Arora, Refrigeration and Air Conditioning, Tata Mcgraw-Hill Publishing Company Limited,Delhi, Third Edition, 2009
[33] AnkitDwivedi, R. S. Mishra," Thermodynamic Analysis of Heat Pipe Using Ammonia, Water and Ethanol with a View to Being Used in Refrigeration" ISSN 2347 - 3258 International Journal of Advance Research and Innovation, Volume 3, Issue 3 (2015) 498-502
[34] AnkitDwivedi, R. S. Mishra, Manjunath K,"Optimization of Vapour Absorption System Using Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET**)**, Volume 5 Issue VIII, August 2017,634-639.

©IJRASET (UGC Approved Journal): All Rights are Reserved

[35] http://hexus.net/tech/news/cooling/81583-fujitsu-thin-loop-heat-pipe-cooler-said-5x-efficient/
[36] https://www.tut.ac.jp/english/schools/faculty/me/post_6.html
[37] AnkitDwivedi, R. S. Mishra," Performance Enhancement of Simple Vapour Absorption System Using Loop Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET)**,** Volume 6 Issue I, January 2018, 865-873.
[38] AnkitDwivedi, R. S. Mishra," Study of Performance Enhancement of Half Effect Vapor Absorption System Using Loop Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET)**,** Volume 6 Issue I, January 2018, 1147-1156.
[39] AnkitDwivedi, R. S. Mishra," Performance Enhancement Analysis of Double Effect Vapor Absorption System Using Loop Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET)**,** Volume 6 Issue I, January 2018,1523-1533.

©IJRASET (UGC Approved Journal): All Rights are Reserved

View publication stats

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/325145143

Study of Performance Enhancement of Half Effect Vapor Absorption System Using Loop Heat Pipes

Article · January 2018

CITATIONS
0

READ
1

2 authors, including:



Ankit Dwivedi
Delhi Technological University
8 PUBLICATIONS **0** CITATIONS

SEE PROFILE

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

Doctoral Research View project

All content following this page was uploaded by Ankit Dwivedi on 15 May 2018.

The user has requested enhancement of the downloaded file.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

Study of Performance Enhancement of Half Effect Vapor Absorption System Using Loop Heat Pipes

Ankit Dwivedi, [1] R S Mishra[2]
[1,2]*Dept. of Mechanical Engineering, Delhi Technological University, Delhi*

Abstract: ***The Half Effect Vapor Absorption Refrigeration System (VARS) has performance lower than the Single Effect VARS due to several reasons. This research aims to enhance the performance of the half effect system by incorporation of a Loop Heat Pipe (LHP) between the high absorber, high generator and condenser (which is actually replaced by the LHP) to avail the intra-cycle heat exchange. The simulations show that COP I and COP II increase by 64 % and 27 % respectively. Also the LHP condenser temperature*** T_{Cond} ***is dependent on the generator temperature*** T_G ***. At higher temperatures of TG the increase in*** COP_{II} ***is more than that in*** COP_{I} ***. Average heat leak from the LHP*** Q_{Leak} ***is around 14.38 kW and average heat utilized due to the LHP*** Q_{Cond} ***is found to be 79.52 kW.***

I. INTRODUCTION

A half-effect vapour absorption system (Fig 3) consists of 2 generators, 2 absorbers, a condenser, an evaporator, 2 pumps, 2 heat exchangers and 3 throttling valves. The half-effect cycle is basically a combination of two single-effect absorption cycles each working at different pressure levels. This system has been developed for relatively low-temperature heat source application. Also the COP of the half-effect system is relatively lower because it rejects more heat than a single-effect cycle. Heat from high temperature external source, transfers to the generator and the absorbers reject heat to the surroundings.

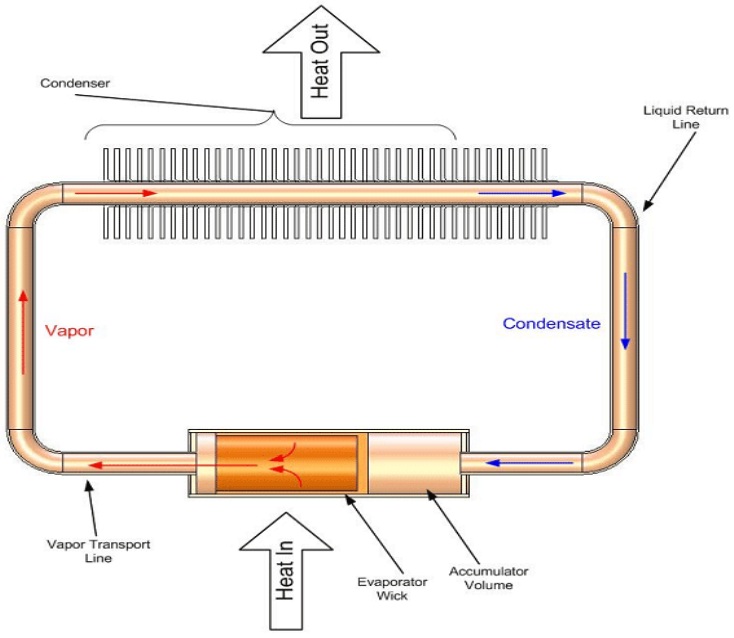


Fig 1: Cyclic process of a Loop heat pipe [36]

Loop heat pipes (LHPs) are two-phase devices used to transfer heat and maintain temperature. It was developed 1980s for spacecrafts uses. Heat in evaporator vaporizes the working fluid at the wick outer surface, then the vapor flows down the system of grooves and through the vapor line towards the condenser. The fluid condenses in the condenser part. A two-phase chamber (compensation chamber) is present at the end of the evaporator which works at a slightly lower temperature (than the evaporator and the condenser). The lower value saturation pressure in the reservoir draws the condensates and liquid return line. Then the fluid flows into a central pipe where it is fed to the wick.

©IJRASET (UGC Approved Journal): All Rights are Reserved

LHPs are self-primed as the volumes of the reservoir, condenser and vapor and liquid lines are controlled so that fluid is always existing to the wick. The reservoir volume and fluid are set to always have fluid in the reservoir even if the condenser and vapor and liquid lines are full of fluid.

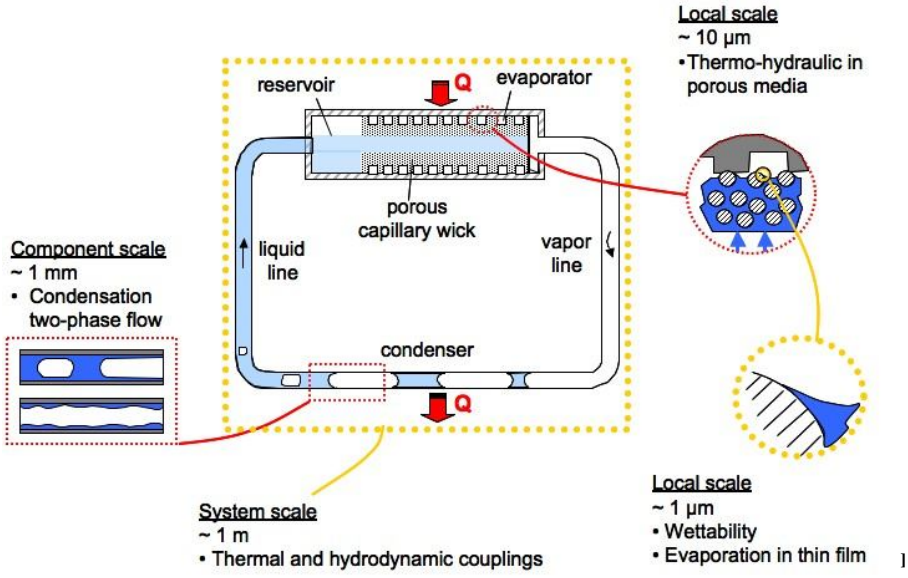


Fig 2: Porous Wick in the LHP.[35]

The LHPs have an inverted wick, and the vapors are located adjacent to the heated surface. These wicks are prepared by power metallurgy. The outer surface of the wick is in contact with the heated surface. Circumferential and axial grooves are required to generate flow channels for the vapors flow which can be machined in the wick, or to the evaporator body.

II. LITERATURE REVIEW

Fabian Korn et al. [2012]performed several vital experiments on heat pipes to establish it to be one of the most effective procedures to transport thermal energy from one point to another, mostly used for cooling[6].Sameer Khandekaret al.[2010]performed experiments on the global thermal performance modeling of Pulsating Heat Pipes (PHPs) requires local, spati-o-temporally coupled, flow and heat transfer information during the characteristic, self-sustained thermally driven oscillating Taylor bubble flow, under different operating conditions[7].JozefHužvár, Patrik Nemecet al. [2007]used heat pipe, observed its basic principles and operating limits. High temperature heat pipes were evaluated for use in energy conversion applications such as fuel cells, gas turbine re-combustors, and Stirling cycle heat sources, with the resurgence of space nuclear power, additional applications include reactor heat removal elements and radiator elements[8].R.Z. Wanget al.[2008]added heat pipes in adsorption water chiller or ice maker initials. His work showed that the adsorption refrigerators are very efficient [10].PrachaYeunyongkul et al.[2009]aimedat experimentally investigating the application of a closed loop oscillating heat pipe (CLOHP) as the condenser for a vapor compression refrigeration system[14].R. Rajashree et al.[1990]went through a numerical analysis of an unsteady, viscous, laminar, incompressible, two dimensional heat and mass transfer, in the vapour gas region of gas loaded circular heat pipe [20]. Da-Wen Sun (1996) performed a detailed thermodynamic analysis of the properties of these binary fluids and expressed in polynomial equations. The performances of three cycles were compared. M.M. Talbi et al. (2000) carried out an exergy analysis on a single-effect absorption refrigeration cycle with lithium-bromide±water as the working Fuid pair. E. Kurem et al.(2001) analyzed the Absorption Heat Pump (AHP) and Absorption Heat Transformers (AHT) using ammonia-water and water-lithium bromide solutions. A fundamental AHP and AHT systems was described and explained the operating sequence. R.D. Misra et al. (2002) applied the therm-o-economic theory is to the economic optimization of a single effect water/LiBrvapour absorption refrigeration system for air-conditioning application. S.A. Adewusi et al (2004). studied the performance of single-stage and two-stage ammonia–water absorption refrigeration systems (ARSs).They calculated entropy generation of each component and the total entropy generation of all the system components as well as COP of the ARSs . S. Arivazhagan et al. (2006) investigated experimentally on the performance of a two-stage half effect vapour absorption cooling system .The prototype is designed for 1 kW cooling capacity using HFC based working fluids (R134a as refrigerant and DMAC as absorbent). RabahGomri et al. (2008) performed exergy analysis of

©IJRASET (UGC Approved Journal): All Rights are Reserved

double effect lithium bromide/water absorption refrigeration system. The system consisted of a second effect generator between the generator and condenser of the single effect absorption refrigeration system, including two solution heat exchangers between the absorber and the two generators. S.C. Kaushik et al. (2009) presented the energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption systems. They developed a computational model for the parametric investigation of the systems. Berhane H. Gebreslassie et al. (2010) performed an exergy analysis, which only considered the unavoidable exergy destruction, conducted for single, double, triple and half effect Water–Lithium bromide absorption cycles. Gulshan Sachdeva et al.(2014) performed an exergy analysis of VAR system using LiBr-H2O as working fluid with the modified Gouy-Stodola approach. Karl Ochsner (2008)et al. (2008) developed a new CO2-heat pipe with high-grade steel corrugated pipe system, which – contrary to other pipe systems permits raw length up to 100 m. They also described the establishment of the heat pump system in general.There are different sorts of the VARS namely half, single, double, triple effects etc. Thorough study of the research papers it can be observed that rigorous thermodynamic and thermo-economic studies have already been performed on almost all of the possible systems. The ground work for further improvements in the systems has also been created. Some other components can be brought in the system and some further enhancements in the performance can be brought in.In the research work LHP can be used as the new component for intra-cycle heat exchange and reduce the requirement of heat input for the cycle to operate. Also replacements of other bulky components can be done by the LHP, which would reduce the cost and size of the system

III. SYSTEMS DESCRIPTION

Figure 3 shows a schematic of Half Effect VARS which is working between 100ºC and 5°C. As the simulation shows the half effect cycle has lower COP $_{I}$& COP $_{II}$ than the single effect cycle, one can apply some modifications to improve the performance of the half effect cycle. The refrigeration capacity of the cycle is set to be 90 kW and the working fluid can be NH_3-H_2O solution. The system consists of 2 generators, absorbers and heat exchangers.

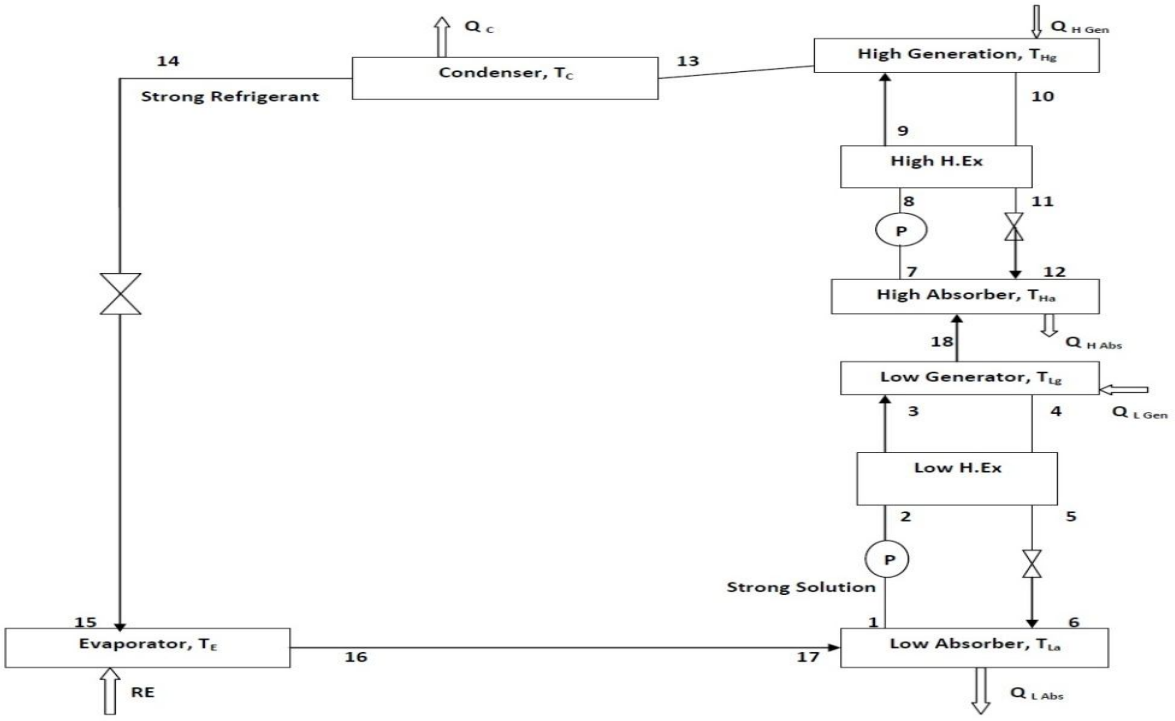


Fig 3: A Half Effect Vapour Absorption System

Figure 4 on the other hand shows the modifications proposed in the above system to have improved performance. There can be a LHP installed between the High Absorber and High Generator removing the heat exchanger and the condenser, thus reducing the size and cost of the system. The LHP acts as the condenser and the heat exchanger altogether. The benefit of high temperature heat transfer will result in decrease in energy and increase in COP $_{II}$. Also it reduces the requirements in the high generator. There can

©IJRASET (UGC Approved Journal): All Rights are Reserved

also be a LHP used between the High absorber and Low Generator similar to the arrangement that will reduce the heat requirements in low generator. The table 1 has the list of the terms used in the research work.

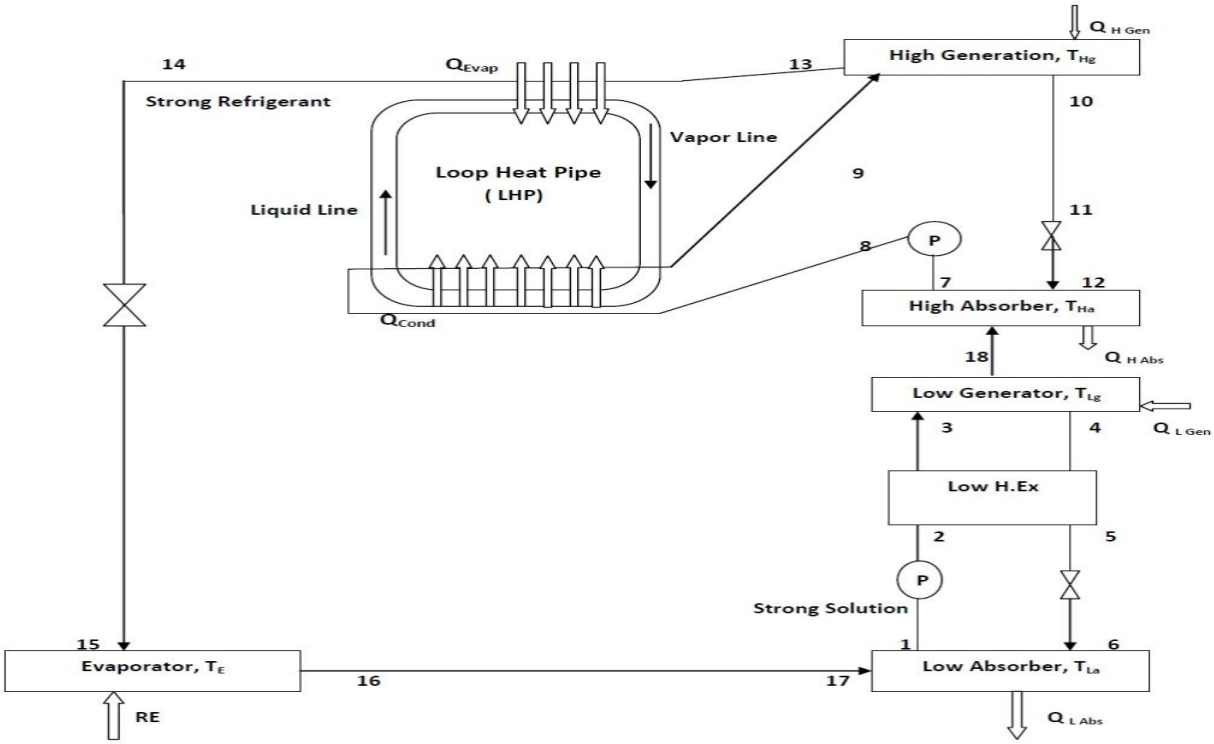


Fig.4: Modified Half Effect Vapour Absorption System

Table 1: Terms Used in Simulation

Terms	Abbreviations
Refrigeration Effect in kW	RE (kW)
Heat rejected in absorber in kW	Q_a(kW)
Heat supplied in generator in kW	Q_g (kW)
Heat rejected in condenser of LHP in kW	Q_{cond}(kW)
Heat absorbed in evaporator of LHP in kW	Q_{eva}(kW)
Absorber Temperature in°C	T_{La},T_{Ha}(°C)
Generator Temperature in °C	T_{Hg}, T_{LG} (°C)
LHP Condenser Temperature in °C	T_c (°C)
Evaporator Temperature in °C	T_E,T_e(°C)
Heat Rejected in Condenser in kW	Q_C (kW)
First Law Coefficient of Performance	COP_I
Second Law Coefficient of Performance	COP_{II}
Heat Leaked from the LHP in kW	Q_{Leak} (kW)
Percentage Improvement in First Law Coefficient of Performance	$\%COP_{I\,imp}$
Percentage Improvement in Second Law Coefficient of Performance	$\%COP_{II\,imp}$
Improvement in First Law Coefficient of	$COP_{I\,imp}$

©IJRASET (UGC Approved Journal): All Rights are Reserved

Performance	
Improvement in Second Law Coefficient of Performance	$COP_{II\ imp}$

IV. RESULTS AND DICSUSSIONS

The simulations show the following results. Fig 5 & Fig 6 show the improvements in COP_{I} and COP_{II} with varying heat utilized in the condenser of LHP (Q_{Cond}) with reference to the performance of the system without modifications. It can be inferred from the Fig 5 that the improvements in COP_{I} is directly proportional to the heat being utilized. The average improved COP_{I} ,that the simulations show, is around 0.69.

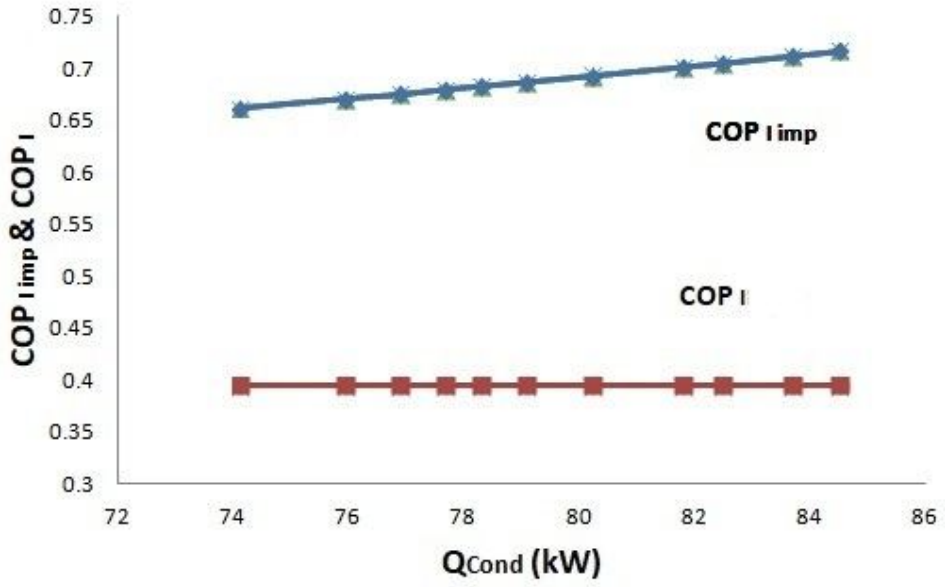


Fig 5: Comparison between COP_{I} and $COP_{I\ imp}$ plotted with Q_{Cond}

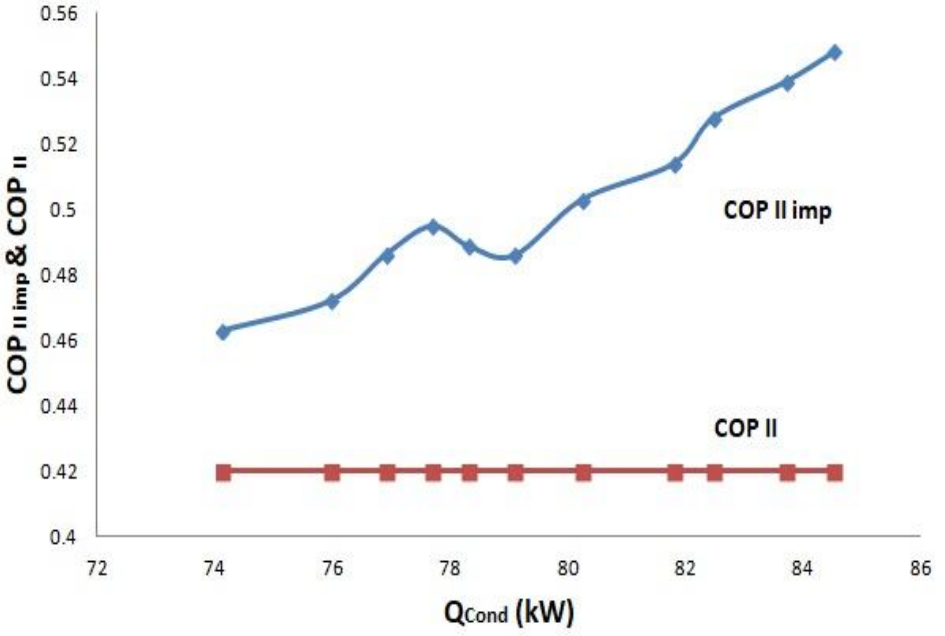


Fig 6: Comparison between COP_{II} and $COP_{II\ imp}$ plotted with Q_{Cond}

©IJRASET (UGC Approved Journal): All Rights are Reserved

The Fig 6 shows that with reuse of heat in the LHP decreases the loss of exergy increasing the COP $_{II}$. The COP $_{II}$ decreases a little but then it rises steadily similar to the COP $_{I}$. Also the heat transfer at high temperatures helps in decreasing the exergy loss. The average enhanced COP $_{II}$ has been calculated to be 0.502 .

Moreover the Fig 7 and Fig 8 show the relative comparison between the COP $_{I}$&COP $_{II}$.Fig 7 shows the comparison between the COP $_{I}$ and COP $_{II}$ with the varying Q_{Cond}. Both rise steadily for the entire range of the Q_{Cond}.

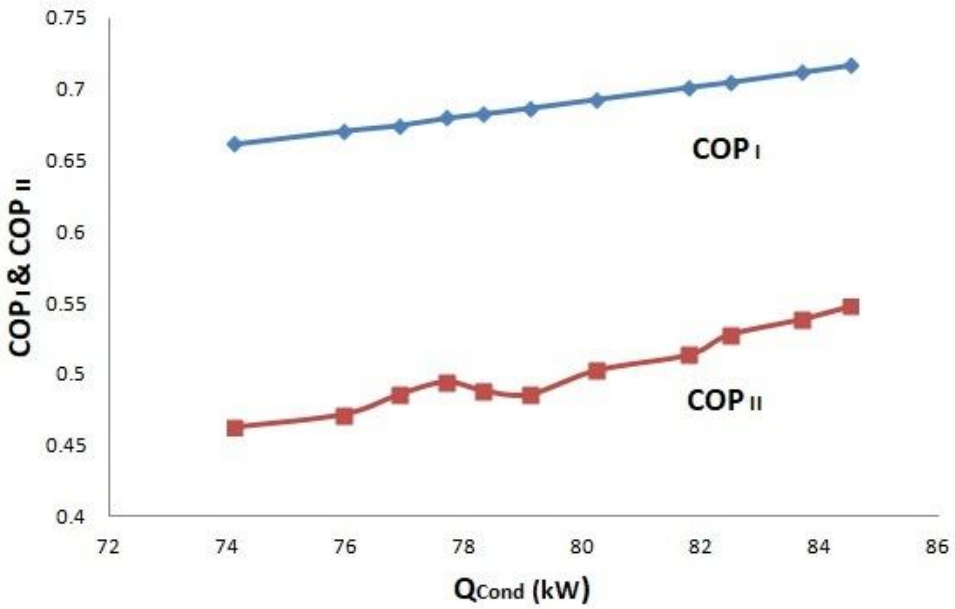


Fig 7: COP $_{I}$ and COP $_{II}$plotted with Q_{Cond}

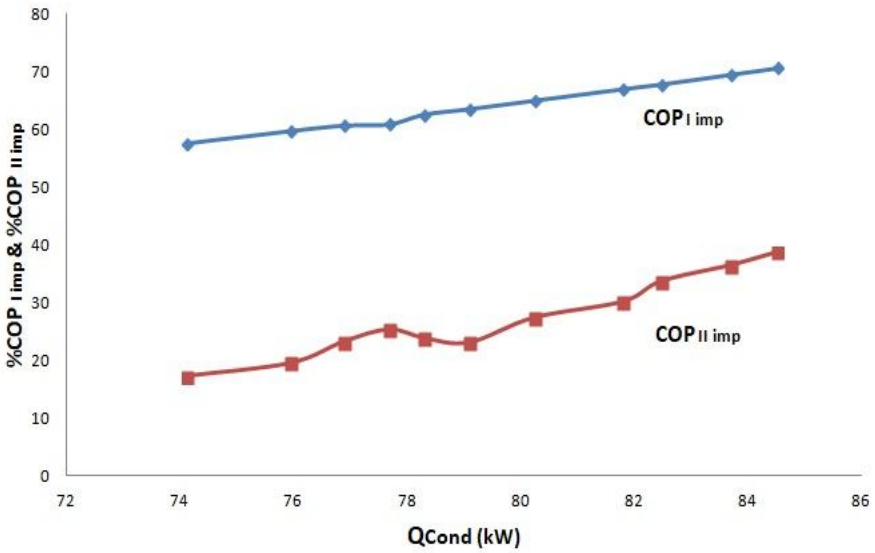


Fig8: Variation of %COP $_{I\,imp}$ and % COP $_{II\,imp}$ plotted with Q_{Cond}

The fig 8 shows the percentage improvement in the COP $_{I}$ and COP $_{II}$ with Q_{Cond}.The average rise in the COP $_{I}$ is 64 %, where as that in COP $_{II}$ is 27 %. In both the figures, COP $_{I}$ and COP $_{II}$ the rise is parallel.

The Fig 9,10,11,12 show the variation of COP $_{I}$ and COP $_{II}$ with T_C and T_G . It can be seen that the relation between the two temperatures is linear. Higher the T_G, higher will be the T_C. With high temperatures of heat exchange, the COP$_{II}$rise is more than the rise in COP $_{I}$. Whereas the rise in COP $_{I}$ becomes flat.

©IJRASET (UGC Approved Journal): All Rights are Reserved

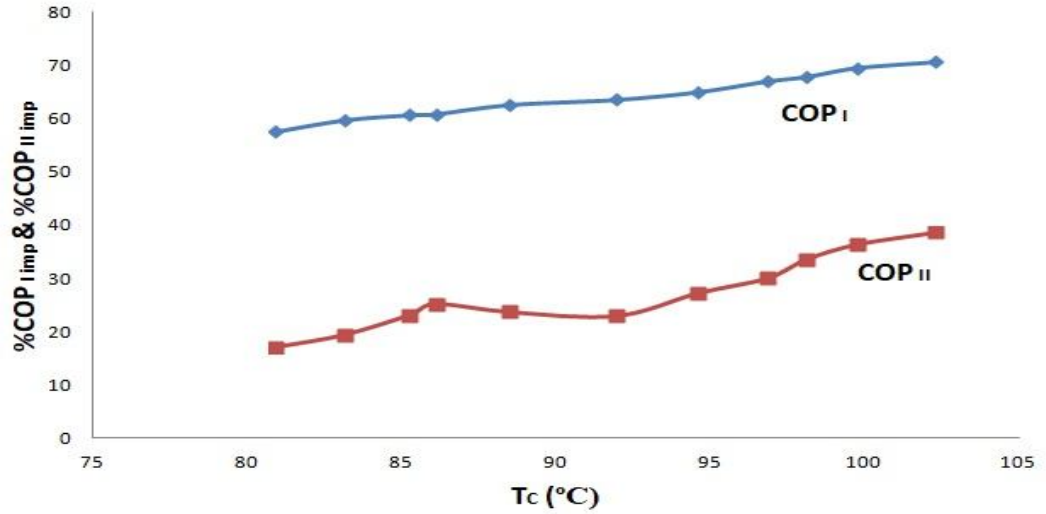


Fig9: Comparison of %age improvements in COP $_{I}$& COP $_{II}$ varying the T_C

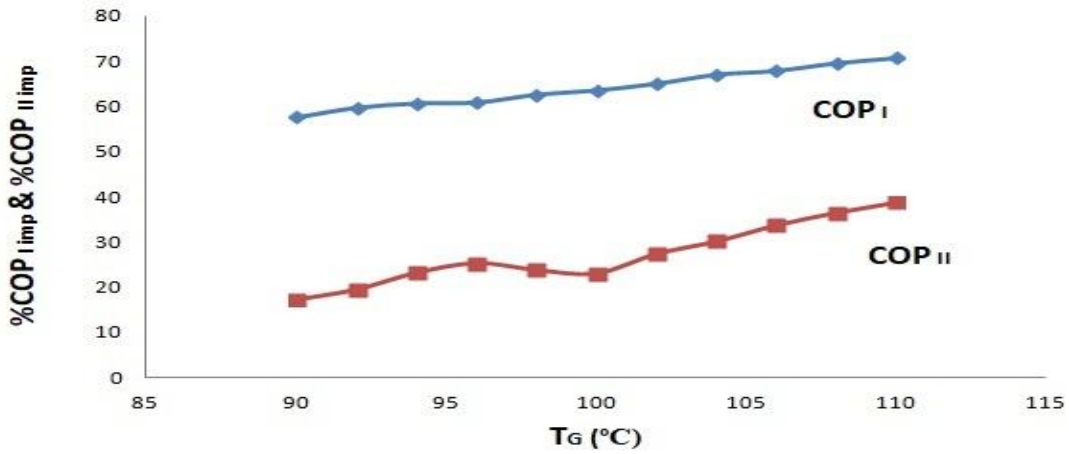


Fig 10: Comparison of %age improvements in COP $_{I}$& COP $_{II}$ varying the T_G

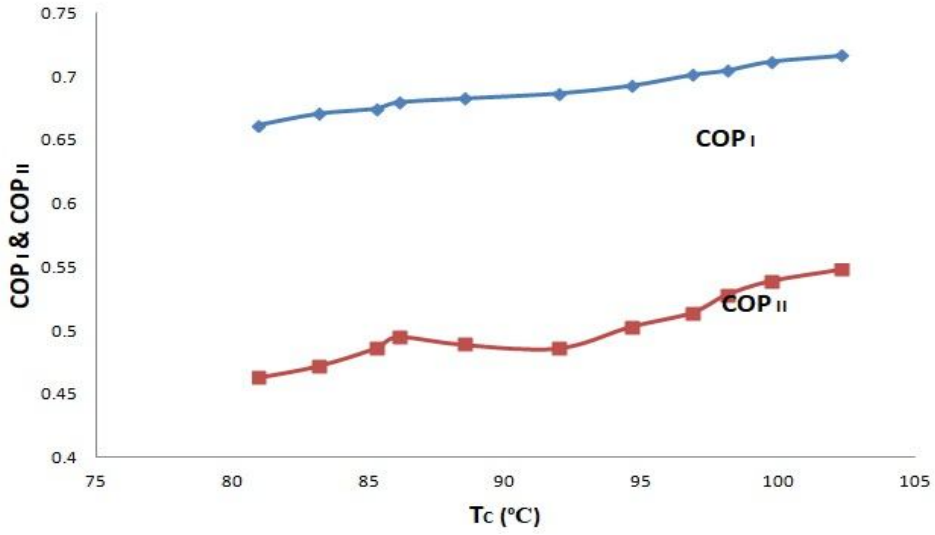


Fig 11: Comparison of COP $_{I}$& COP $_{II}$ varying the T_C

©IJRASET (UGC Approved Journal): All Rights are Reserved

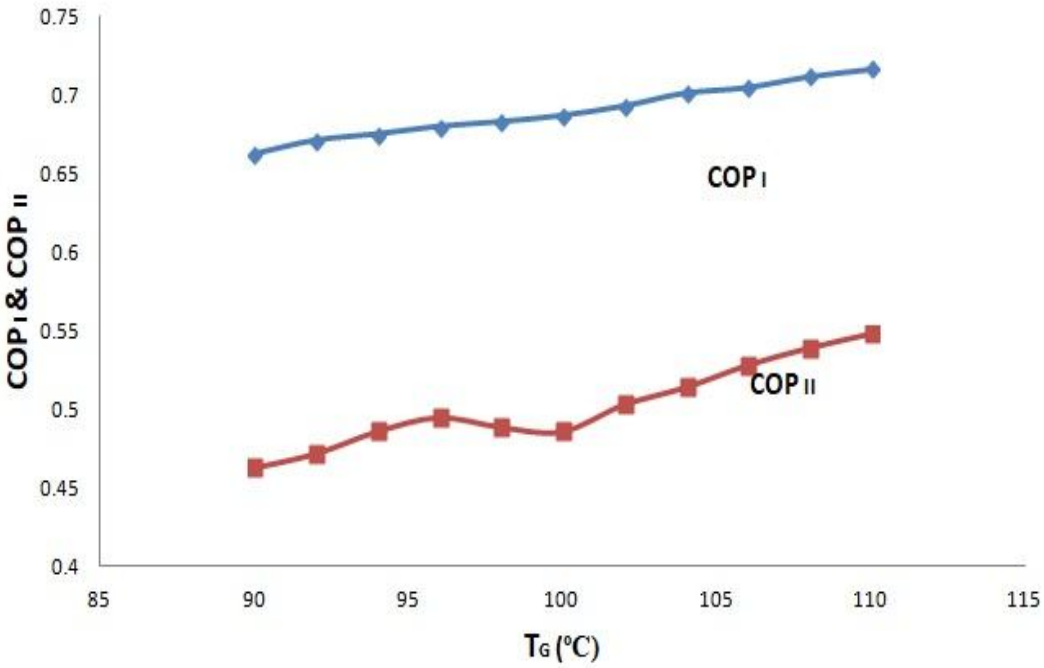


Fig 12: Comparison of COP $_{I}$& COP $_{II}$ varying the T_C

The Fig 13 & Fig 14 show the values of Q_{Cond}&$Q_{Leak}$with respect T_C and T_G. With increase in the temperatures the heat leaked from the LHP is found to reduce, whereas the heat reusable increases.

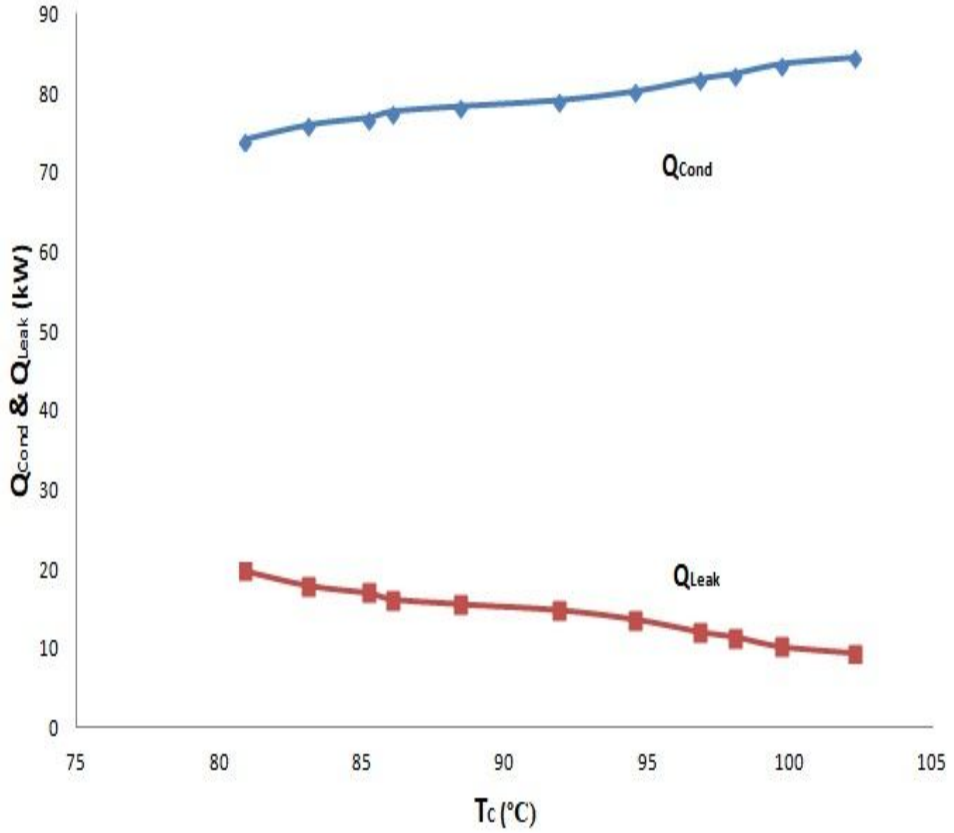


Fig 13: Comparison of Q_{Leak}&Q_{Cond}varying the T_C

©IJRASET (UGC Approved Journal): All Rights are Reserved

The T_C is a dependent parameter on T_G directly. As the T_G increases the Tc also increases the heat to be supplied and the maximum temperature up to which it can be transferred also increases. The average $Q_{Cond}$is observed to be 79.52 kW.

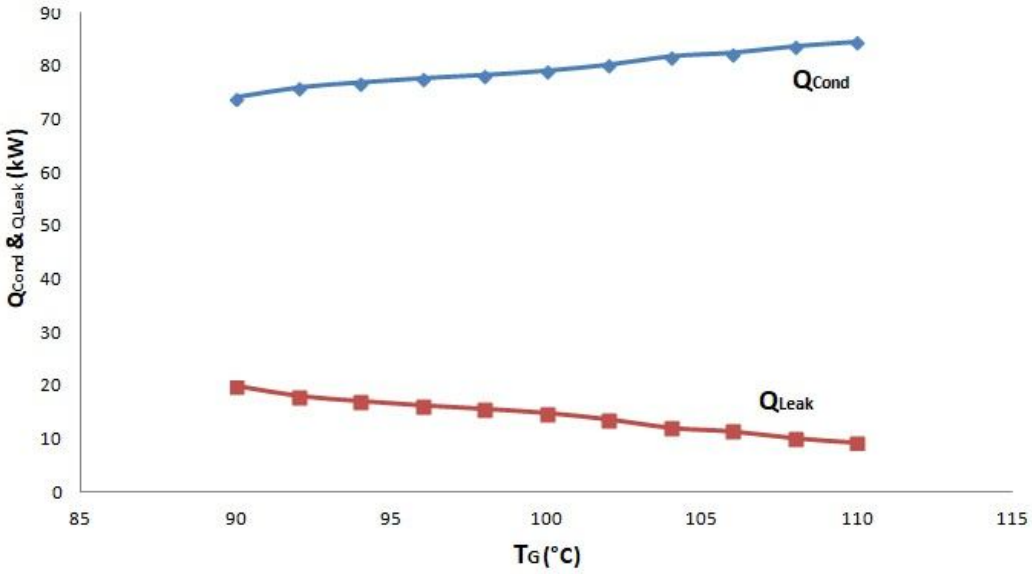


Fig 14: Comparison of Q_{Leak}&Q_{Cond}varying the T_G

$Q_{Leak}$is found to decrease as the operating temperature increases. The operating temperature increases the heat transfer coefficient hence more heat is transferred in the evaporator of the LHP, reducing the Q_{Leak}. The average Q_{Leak} is calculated to be 14.38 kW.

V. CONCLUSIONS

Through the simulations followings can be the conclusions for the Half Effect VARS:

A. COP_I& COP_{II} increase with the Q_{Cond} and the average values for the modified value are 0.69 & 0.502 respectively.
B. The average rise in the COP_I is 64 %, where as that in COP_{II} is 27 %.
C. Higher the T_G, higher will be the T_C. With high temperatures of heat exchange, the COP_{II} rise is more than the rise in COP_I. Whereas the rise in COP_I becomes flat.
D. With increase in the temperatures the heat leaked from the LHP is found to reduce. Average $Q_{Leak}$is around 14.38 kW.
E. The operating temperature increases the heat transfer coefficient hence more heat is transferred in the evaporator of the LHP having the average Q_{Cond}as 79.52 kW.
F. The exergetic losses can be reduced, along with the reduction in size of the system, making the system flexible.

REFERENCES

[1] Saeed. Sedigh , Hamid. Saffari ,"Thermodynamic analysis of single effect and half effect absorption refrigeration system" International Journal of Energy & Technology Vol.25 (2011) 1-9.
[2] S. Arivazhagan , R. Saravanan , S. Renganarayanan ," Experimental studies on HFC based two-stage half effect vapour absorption cooling system" Applied Thermal Engineering Vol. 26 (2006) 1455–1462.
[3] GulshanSachdeva, Ram Bilash," Thermodynamic Analysis of a Vapor Absorption System Using Modified Gouy-Stodola Equation" International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:8, No:12, 2014.
[4] I. Horuz," A comparison between Ammonia-water and Water-Lithium Bromide solutions in Vapor Absorption Refrigeration Systems" Int. Comm. Heat Mass Transfer, Vol. 25, No. 5, pp. 711-721, 1998.
[5] Abdul Khaliq, and Rajesh Kumar," Exergy analysis of double effect vapor absorption refrigeration system" Int. J. Energy Res. 2008; Vol.32:161–174.
[6] S.C. Kaushika, AkhileshArora," Energy and exergy analysis of single effect and series flow double effect water–lithium bromide absorption refrigeration systems" international journal of refrigeration Vol.32 (2009) 1247 – 1258.
[7] RabahGomri , RiadHakimi," Second law analysis of double effect vapour absorption cooler system" Energy Conversion and Management Vol.49 (2008) 3343–3348.
[8] S.A. Adewusi, Syed M. Zubair," Second law based thermodynamic analysis of ammonia–water absorption systems" Energy Conversion and Management Vol. 45 (2004) 2355–2369.

©IJRASET (UGC Approved Journal): All Rights are Reserved

[9] RabahGomri," Second law comparison of single effect and double effect vapour absorption refrigeration systems" Energy Conversion and Management Vol. 50 (2009) 1279–1287.
[10] R.D. Misra, P.K. Sahoo, S. Sahoo, A. Gupta,"Thermoeconomic optimization of a single effect water/LiBrvapour absorption refrigeration system" International Journal of Refrigeration Vol.26 (2003) 158–169.
[10] M. Belghazi, A. Bontemps, C. Marvillet," Experimental study and modelling of heat transfer during condensation of pure fluid and binary mixture on a bundle of horizontal finned tubes" International Journal of Refrigeration Vol.26 (2003) 214–223.
[11] M.M. Talbi, B. Agnew," Exergy analysis: an absorption refrigerator using lithium bromide and water as the working Fuids" Applied Thermal Engineering Vol. 20 (2000) 619-630.
[12] E. Kurem," A comparison between Ammonia-water and Water-Lithium Bromide solutions in vapour absorption heat transformers "In. Comm. Heat Mass Transfer; Vol. 28, No. 3, pp. 421-438, 2001.
[13] Da-Wen Sun," Comparison of the performances and NH3-H20, NH3-LiNO3 and NH3-NaSCN Vapor Absorption Refrigeration Systems" Energy Convers. MgmtVol. 39, No. 5/6, pp. 357-368, 1998.
[14] Yu.F. Maydanik," Review Loop heat pipes" Applied Thermal Engineering Vol.25 (2005) 635–657.
[15] Randeep Singh, AliakbarAkbarzadeh , MasatakaMochizuk," Operational characteristics of a miniature loop heat pipe with flat evaporator" International Journal of Thermal Sciences Vol.47 (2008) 1504–1515.
[16] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, C.J. Chen," Performance study of a high efficient multifunction heat pipe type adsorption ice making system with novel mass and heat recovery processes" International Journal of Thermal Sciences Vol.46 (2007) 1267–1274
[17] Yuan-Ching Chiang , Wen-Cheng Kuo , Chia-Che Ho , Jen-JieChieh," Experimental study on thermal performances of heat pipes for air-conditioning systems influenced by magnetic nanofluids, external fields, and micro wicks" International Journal of Refrigeration Vol.43 (2014) 62 -70.
[18] T.X. Li, R.Z. Wang, L.W. Wang, Z.S. Lu," Experimental investigation of an innovative dual-mode chemisorption refrigeration system based on multifunction heat pipes" International Journal of Refrigeration 3 1 (2 0 0 8) 1 1 0 4 – 1 1 1 2.
[19] L. GarousiFarshi a,*, C.A. Infante Ferreira b, S.M.S. Mahmoudi a, M.A. Rosen," First and second law analysis of ammonia/salt absorption refrigeration systems" International Journal of Refrigeration Vol. 4 0 (2 0 1 4) 1 1 1-1 2 1.
[20] T.X. Li, R.Z. Wang , L.W. Wang, Z.S. Lu, J.Y. Wu," Influence of mass recovery on the performance of a heat pipe type ammonia sorption refrigeration system using CaCl2/activated carbon as compound adsorbent" Applied Thermal Engineering Vol. 28 (2008) 1638–1646.
[21] Z.S. Lu , L.W. Wang, R.Z. Wang," Experimental analysis of an adsorption refrigerator with mass and heat-pipe heat recovery process" Energy Conversion and Management Vol. 53 (2012) 291–297.
[22] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences Vol.50 (2011) 592-598.
[23] Basant K. Agrawal*, Munawar N. Karimi," Thermodynamic performance assessment of a novel waste heat based triple effect refrigeration cycle" International Journal of Refrigeration Vol. 3 5 (2 0 1 2) 1 6 4 7-1 6 5 6.
[24] Behrooz M. Ziapour*, Mohsen Tavakoli," Performance study on a diffusion absorption refrigeration heat pipe cycle" International Journal of Thermal Sciences 50 (2011) 592-598.
[25] T.S. Jadhav , M.M. Lele," Theoretical energy saving analysis of air conditioning system using heat pipe heat exchanger for Indian climatic zones" Engineering Science and Technology, an International Journal Vol. 18 (2015) 669-673.
[26] Chengchu Yan , Wenxing Shi , Xianting Li , Shengwei Wang," A seasonal cold storage system based on separate type heat pipe for sustainable building cooling" Renewable Energy Vol.85 (2016) 880-889.
[27] Matthias H. Buschmann," Nanofluids in thermosyphons and heat pipes: Overview of recent experiments and modelling approaches" International Journal of Thermal Sciences Vol.72 (2013) 1-17.
[28] P.D. Dunn, D.A. Reay, Heat Pipes, Pergamon Press, Oxford, 1993.
[29] D.Reay, Heat Pipes-Theory, Design and Applications, Butterworth-Heinemann,Oxford, Fifth edition 2006.
[30] Korn, F., "Heat Pipes and its Applications" Project Report 2008 MVK160 Heat and Mass Transport May 07, 2008, Lund, Sweden.
[31] Rajashree, R., Rao, K.S., "**A** Numerical Study of the Performance of Heat Pipe" Indian Journal of Pure and Applied Mathematics, 21 (1): 95-108, January 1990.
[32] C.P. Arora, Refrigeration and Air Conditioning, Tata Mcgraw-Hill Publishing Company Limited,Delhi, Third Edition, 2009.
[33] AnkitDwivedi, R. S. Mishra," Thermodynamic Analysis of Heat Pipe Using Ammonia, Water and Ethanol with a View to Being Used in Refrigeration" ISSN 2347 - 3258 International Journal of Advance Research and Innovation, Volume 3, Issue 3 (2015) 498-502.
[34] AnkitDwivedi, R. S. Mishra, Manjunath K,"Optimization of Vapour Absorption System Using Heat Pipes" ISSN: 2321-9653 International Journal for Research in Applied Science & Engineering Technology (IJRASET**),** Volume 5 Issue VIII, August 2017,634-639.
[35]http://iusti.univ-provence.fr/document.php?pagendx=11288&project=iusti [36]https://www.1-act.com/innovations/heat-pipes/heat-pipe-loops/

©IJRASET (UGC Approved Journal): All Rights are Reserved

View publication stats

National Conference on Recent Advances in Computer Science and IT (NCRACIT)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 1 | ISSN : 2456-3307

Subspace-Based Adaptation of Detectors for Video

Sangeeta Bhan[1], Sunil Dalal[2], Pankaj Choudhary[3]

[1]Computer Science and Engineering, Delhi Technological University, Delhi, India
[23]Department of Information Technology BGSB University, Rajouri, J&K, India
sangeetabhan7@gmail.com[1], sunildalal57@gmail.com[2], pschoudhary@bgsbu.ac.in[3]

ABSTRACT

Object detection in videos has always been a challenging problem to work with. Detection of a particular class object plays an important role in many real-world applications. Since the domain of source and target video vary significantly, classifier being trained on source video does not give expected results on the target video. Thus, domain adaptation techniques are used, one of which is Subspace Based Adaptation. In this technique, first, we compute both source and target subspace from the features collected. Since we do not have target data directly, we use different ways to get data from the target video. Compute subspace after collecting the data from both source and target videos. Eigen vectors describe this generated source and target subspaces.

Keywords: Subspace, Detector, Adaptation, Histogram.

I. INTRODUCTION

In detection the object of the particular class, it is typically presumed that source and target data have the same distribution. However, it is not true in real world applications. We are training our simple detector model from the source video with the help of its annotation file. Initially, the problem is to collect the negatively labeled data from video. We have used Hard *Negative Mining* to collect this negatively labeled data. Next, we extract the features of these data using *Histogram of Oriented Gradients*. With the help of *Background Subtraction using MoG*, data samples were collected from the target video for subspace calculation. PCA is used to find the best Eigen vectors for both source and target dataset. Once we have calculated the subspaces using PCA, transforms the source subspace coordinate system into the target aligned source subspace coordinate system by aligning the source basis vectors with the target ones. Generate training data by mapping source data into this target aligned source subspace. Now, we train our linear SVM model using this new labeled data. Iteratively use online samples or bounding boxes detected on target video to generate new detector model.

II. RELATED WORK AND BACKGROUND THEORY

Over the past several years, many different have been pro-posed to detect objects of a particular class in videos and images. The main problem in object detection is varying domains of source and target data. So the main issue is to find out the relationship between these two domains. Domain adaptation is a widely used technique in computer vision and language processing. A classical strategy related to our work consists of learning a new domain-invariant feature representation by a new projection

space. Work on PCA based domain adaptation [1] has been done to minimize the marginal distributions between the different domains. Other strategies have been explored on image datasets as well, such as using metric learning approaches [2,3] or canonical correlation method [4] over different views of the data to find a coupled source-target subspace where one assumes the existence of a performing linear classifier on the two domains. Boosting-based [5] approaches are used for detection of objects in image or video. First, they train the initial model on some small labeled datasets and then use this model to collect a larger labeled dataset then train the new model iteratively by using this new dataset. Object detection and classification is the area which has received much attention in the research of computer vision and pattern recognition recently.

III. METHODOLOGY

The very first step in novel algorithm for detection is to collect all the labeled data from source video and unlabeled data from target video [Algorithm 1]. Next calculate subspace using PCA by selecting the top d eigenvectors, to train the detector [Algorithm 2]. We learn transformation matrix by minimizing the Bregman matrix divergence [1].

At the time of testing, iteratively learn new detector by the help of detected bounding boxes [Algorithm 3].

Algorithm 1 Data Collection

1: **Given:** Source Video V_S, Target Video V_T , Source Annotation File A_S, Object Class C
2: **Init:** Positive labeled data P = {}, Negative labeled data N = {}, Number of random samples R = 100, δ= 100
3: **for** each frame $i \in V_S$ **do**
4: $H_n \leftarrow 0$
5: $P \leftarrow$ Extract & Compute features of object ∈ C from i using A_S
6: $N \leftarrow$ Extract & Compute features of object ∉ C from i using As
7: $N \leftarrow$ Extract & Compute features of δ random samples from i
8: simpleSVM ← trainSVM(P,N)
9: boundingBoxes←runsimpleSVM(i)
10: N ← FP(boundingBoxes) using A_S
11: δ←R-CountofFP(boundingBoxes)
12: end for
13: S ← merge(P ,N)
14: for each frame $i \in V_T$ **do**
15: matrixM←ForegroundMaskUsingMoG(i)
16: matrixM ← Filters(matrixM)
17: AllContours←detectContours(matrixM)
18: **for** each contour $k \in$ AllContours **do**
19: T ← Extract & Compute features of k . T is Target Data
20: end for
21: **end for**

Sets. Since a set of all the possible window sized images in the form of patches at different scales and locations from a frame provides a nearly endless supply of negative samples. Training with all the available data is considered impractical.

Algorithm 2 Generate Subspace & Train Detector

1: **Given:** Source data S, Target data T , Subspace dimension d
2: $X_S \leftarrow$ runPCA(S,d)
3: $X_T \leftarrow$ runPCA(T,d)
4: M ← XS' XT ⊳ XS' is transpose of XS & M is Transformation Matrix
5: Xa ← XSM ⊳Xa is New Coordinate System
6: trainData ← SX
7: detectorModel←LinearSVM(trainData,Labels)

Algorithm 3 Use of Online samples to Generate Detector

1: **Given**: Source data S, Target data T, Subspace dimension d, Initial detector-Model, Positive data P, Negative data N, Target Video V_T
2: **for** each frame $i \in V_T$ **do**
3: boundingBoxes← rundetectorModel(i)
4: **for** each bounding box B ∈ boundingBoxes do

5: Use Sliding Window Method for B in i to Categorize into TP & FP

6: Add B into S with Label

7: **end for**

8: Generate subspace and Calculate trainData using Algorithm 2

9: detectorModel← trainSvm(newtrainData)

10: **end for**

IV. EXPERIMENTS AND RESULTS

The proposed algorithm was implemented in C++11 with Open CV 3.0. We used VIRAT video dataset to train and test our detector models. These videos are captured by stationary HD cameras and each video has its own annotation file which carry information about bounding box of objects in each frame of the video. . The virat_s_050000_03 sequence is used as training video and VIRAT_S_040103_02 sequence as test video. These sequences have 1607 and 2392 frames respectively, each of size 1920 X 1080. All extracted objects from source and target video, are scaled to the size of 64x64. Features are extracted and computed using HOG, and thus, descriptor vectors are 1764-dimensional. As a detector model, we used Linear SVM with C=0.01.

Experiments are conducted extensively for different methods using adaptation and online samples on videos. Inferences obtained from results are analyzed in detail. For each method, average precision over all the frames is plotted against the number of top k bounding boxes in precision @K.

Here in this method, we train the model on HOG features, extracted from source or target video. Some of the parameters are to be set in HOG for the best results. Figure 1 shows the relation between precision value and K where K indicates the top bounding boxes. The mean precision and recall over target video using this method are 51 and 19 percent respectively.

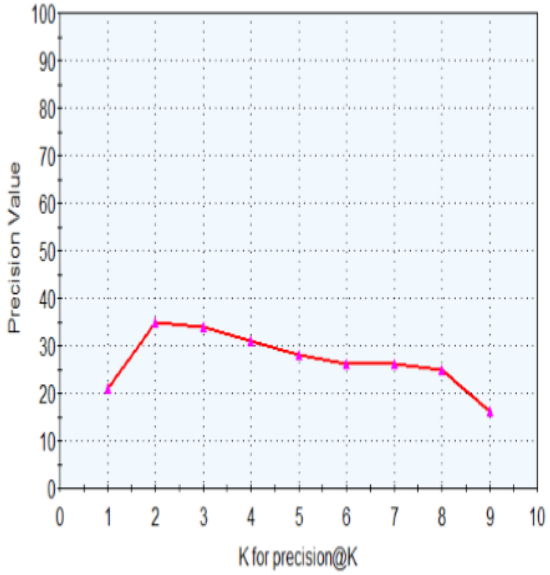


Figure 1.

Adaptation is based on the subspaces generated from source and target video. Here in this section, we take 100(it may vary) samples from each of the frame target video and then generate target subspace using this random samples.

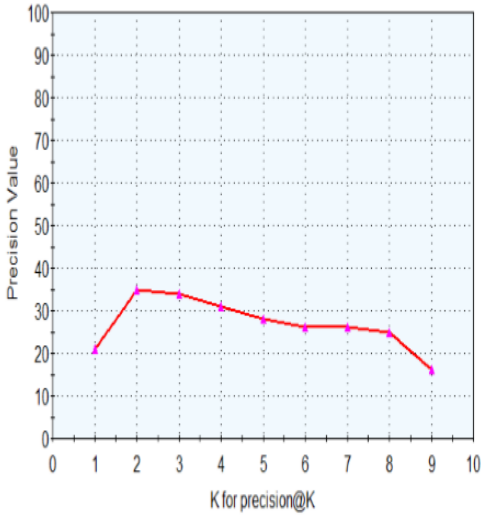


Figure 2.

Overall accuracy is not affected much, since it may happen that no target class object appears in the subspace due to randomness. The mean precision graph for this method is shown in Figure 2.

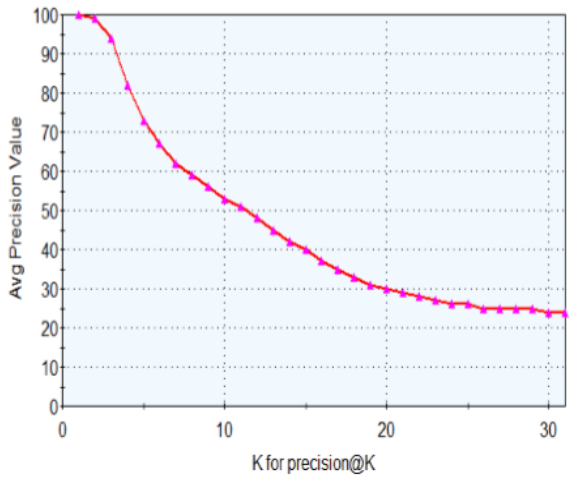


Figure 3.

Given Figure 3 shows that nearly all top bounding boxes belong to true positives. This method shows a slight improvement in accuracy (Figure 5.3) but still not as expected. The mean precision and recall for this method is 28 and 53 percent respectively.

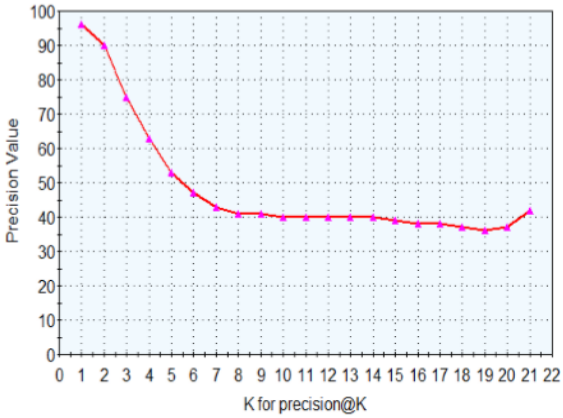


Figure 4.

Here we go with the Mixture of Gaussians (MoG), for subtracting background from a frame to get foreground mask. It extracts all the moving objects in the target video. This data generates a good target subspace, resulting in higher accuracy. The mean precision and recall in this method are 42 and 41 respectively which are greater than the values achieved in all the previous methods. For this method, the plot of precision is shown in Figure 4.

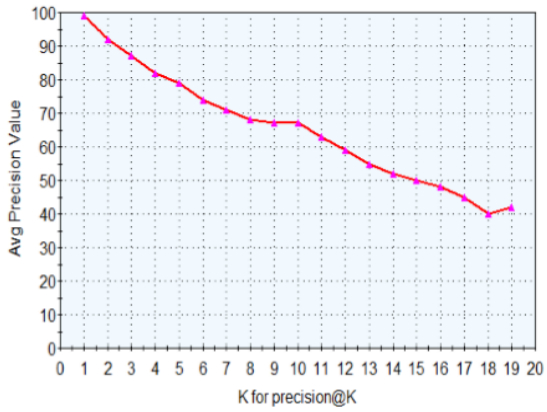


Figure 5.

The mean precision and recall for this method is 56 and 60 percent respectively, which is extremely well as compared to all the previous methods. Precision @K graph for this method is shown in Figure 5.

Comparison of all Methods

The very first method we used is based on simply HOG and SVM (Method1). Here we extract the features of training data (From source video) without any use of target data and train the Linear SVM model on these features. Now use this model directly to detect objects in target video. So, this is the simplest method which we used in detection and as we see in the graph (Figure 4.6), results obtained are not so good. As video dataset has different domains, we align source subspace to target subspace. Choosing the random samples to calculate subspace of target video does not affect much in accuracy (Method2). So, we used background subtraction on target video, using MoG, to get foreground mask (Method4). This foreground mask is further used to calculate target subspace for adaptation. Now here we saw an improvement in accuracy. Iterative training using weakly labeled online samples gives fairly well results, as we see in Figure 6.

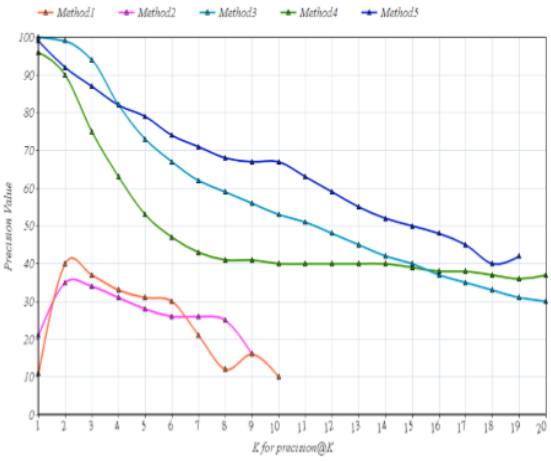


Figure 6.

The graphs represents that use of adaptation in videos, using background subtraction to collect target data and learning from online samples, improves the accuracy of object detection. So we merge both of the method (Method4 and Method3) which gives extremely well results. The bar graph (Figure 7) shows mean precision and recall value for all the methods and both looks fairly well in the Method5.

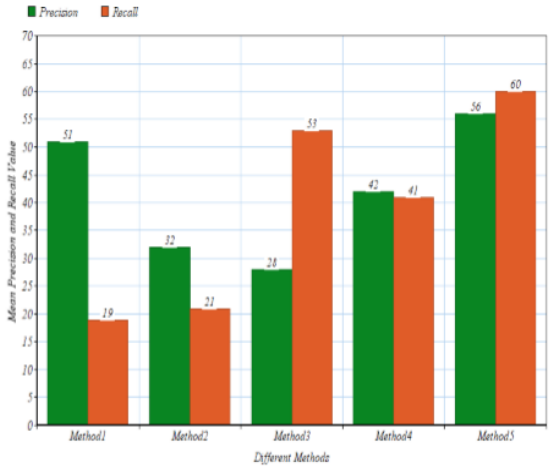


Figure 7.

V.CONCLUSION

We introduced a novel method which uses online samples with subspace based adaptation in videos. Here in this method, first we created subspaces for both source and target video domains and then learned a linear mapping that aligns the source subspace with the target subspace. This allows us to build a detector model on source data in target based space which can be applied to the target video. Different methods to generate detectors with different parameters like block Size, cell Size, n bins (Number of bins) in HOG, C-value, ϵ in SVM and d-value to generate subspace using PCA, and were compared simultaneously for VIRAT video datasets having different domains. We have collected data from target video using several different methods from each of the target video frame and detected objects, as future work, we intend to improve performance of object detection in videos by using some other methods. The novel method introduced by us by creating subspaces for both source and target video domains and builds a detector model on source data in target based space can be improved by using some other improved training methods.

VI.REFERENCES

[1] M. Sebban B. Fernando1, A. Habrard2 and T. Tuytelaars1. Unsupervised visual domain adaptation using subspace alignment. In ICCV, 2013.

[2] M. Fritz K. Saenko, B. Kulis and T. Darrell. Adapting visual category models to new domains. In Computer Vision ECCV, 6314:213{226, 2010.

[3] K. Saenko B. Kulis and T. Darrell. What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In CVPR, pages 1785{1792, 2011.

[4] D. Foster J. Blitzer and S. Kakade. Domain adaptation with coupled subspaces.

[5] S. Ali O. Javed and M. Shah. Online detection and classi cation of moving ob-jects using progressively improving detectors. Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, pages 1063{6919, 2015.

International Journal of Emerging Technologies in Engineering Research (IJETER)
Volume 6, Issue 5, May (2018) www.ijeter.everscience.org

Surface Composite Fabrication through TIG arc Process: A Review

N. Yuvaraj
Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

Abstract – Surface modification of Materials are essential in industrial applications in order to improve tribological, mechanical and corrosive properties. TIG arc surfacing is an alternative technique for modifying the surface layer and fabrication of surface composites. TIG arc surfacing method improves the surface in an efficient, user friendly and economical manner. This review article narrates the current status of the surface composite fabrication through TIG arc method.

Keywords: Surface modification, surface composite, TIG arc Process, Grain size, Hardness

1. INTRODUCTION

Surface Engineering (SE) is a multidisciplinary activity which deals with the properties of the surfaces of the solid matter. The function and serviceability of the engineering components can be improved through SE. It involves in improving the properties of the surface layer of the materials in order to minimize the degradation over the period [1]. Surface engineering can be defined as "treatment of the surface and near-surface regions of material to allow the surface to perform functions that are distinct from those functions demanded from the bulk of the material" [2]. The surface of the component has to be manufactured in a desirable way to meet the suitable service life and robust to the environment.

Surface modification process changes the properties of the surface at the certain depth level of the substrate. Physical Vapor Deposition (PVD), and Chemical Vapor Deposition (CVD) are common techniques used to deposit a thin overlay (coatings) on the substrate in the range of 1-10 µm and small interaction takes place on the substrate material [3]. In the thermal spray process a significant amount of interaction takes place on the substrate. In the carburizing or nitriding process without adding any new material on the top of the substrate, only surface modification takes place for enhancement of surface properties.

TIG arc surface method is used to modify the surface layer and for producing a thick surface composite layer; the layer thickness can range from hundreds of micrometers to several millimeters. The techniques illustrated in Fig. 1 involve the classification of surface modification processes by the depth and place of interaction and depth on the substrate.

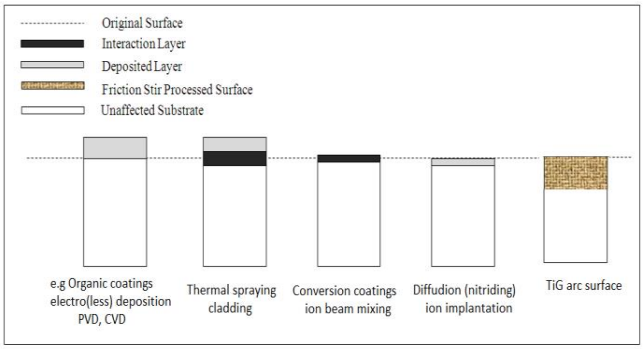


Fig. 1 Classification of surface modification processes by the depth and place of interaction and depth on the substrate [4] [Adopted and Modified]

2. SURFACE COMPOSITE FABRICATION BY TIG ARC PROCESS

The surface property of the material is very important to achieve longer life of mechanical components. To obtain the surface modification, the ceramic particles were applied on the surface of the matrix material by various techniques for enhancing the surface properties of the material. The most common methods are coating, plasma spraying, cladding, laser melt treatment, electron beam irradiation, etc. [5-6]. In these techniques, processing of surface composites is based on liquid-phase processing at high temperature. There will be an interfacial reaction between the matrix and reinforcement and the possibilities of formation of some detrimental phases are high. The ceramic particles were applied to the matrix surface and processed by TIG arc, electron beam irradiation, laser cladding techniques for enhancing the surface properties of the material.

S.H. Choo et al. (1999) [7] have deposited TiC particles on the plain carbon steel plate using high energy electron beam irradiation technique for producing the surface composite. This composite exhibited 3 to 4 times higher hardness than the base alloy due to the presence of TiC hard particles. These processes create an excellent bond between the substrate and the composite. L. Dubourg et al. (2002) [8] have modified the surface by addition of copper and iron powders on the aluminum surface by laser cladding. The hardness of Al/38Cu-11Fe(wt%) hardness was 550 $Hv_{0.2}$ whereas the

hardness of the non-treated sample hardness was 20 $Hv_{0.2}$. Presence of intermetallic compounds, grain refinement and hardened surface of the aluminum matrix improve the hardness of the composite.

S. Jiguo et al. (2007) [9] have used the TIG arc source with Ni and aluminum powder preplaced on the Q235 steel plate to prepare the surface with Ni-Al intermetallics. The different current values and surfacing speed were chosen to make the beads of intermetallics. The combination of high current value of 120 A and speed of 1.5 mm/s the intergranular phase of α-Fe is formed. In the low current value of 80 A and speed of 0.9 mm/s the intergranular phase of γ-(Fe, Ni) with the large amount of Ni-Al is formed.

S. Islak et al. (2012) [10] packed with the mixture of FeW and B_4C particles in different particles ratio's (10%, 20% and 40% ratio of B_4C in FeW/B_4C mixture) in the 8mm diameter and 1.5mm deep holes in the AISI 1060 substrate. TIG arc source was used to produce the surface composite on the steel substrate. The coating thickness of 2000-2400µm was produced on the steel surface. Microhardness of the composite was 4-5 times higher than the base material. The maximum hardness value of 1095Hv was obtained at 40% ratio of B_4C powder in mixture. The base material hardness was around 225Hv. The enhancement of the hardness in the surface coating was due to the formation of carbide and boride [Fe_3(C, B) and Fe_{23}(C, $B)_6$] in the microstructure.

Similarly, S. Mridha et al. (2001) [11] mixed 53% titanium powder and 47% aluminum powder with organic binders and preplaced on the pure titanium substrate and maintained the thickness around 0.6 to 1mm on the surface. The TIG arc source operating current value of 50 to 100A and at the different speeds range of 2, 3 & 4 mm/s was selected for melting the powder to form an intermetallic coating on the substrate. The maximum melt depth of 2mm was obtained when glazed at 100A TIG arc current and speed of 2 mm/s. The energy density of the torch is the main reason for the deeper depth formation of the alloyed layer. The XRD peaks showed that the low energy density (60A) TIG arc produced TiAl structure and the high energy density (110A) TIG arc produced Ti_3Al and TiAl structure. The hardness and wear resistance of the coated surface was improved about 3-4 times than base material. In another study conducted on AISI 8620 steel, SiC/C particles diffused on the surface by TIG arc method. The hardness and wear behavior of the modified surface was significantly improved [12].

Similarly, W. Xinhong et al. (2006) [13] deposited a mixture of Fe based self-fluxing FeCrBSi powder, graphite, and ferrotitanium powder on AISI 1045 steel substrate and

developed a in-situ method coating to produce TiC particles reinforced Fe-based surface composite through TIG process. From the graphite and ferrotitanium powder, the TiC particles were synthesized in the range of 3-5µm and dispersed in the matrix in a uniform manner. The hardness and wear resistance of the composite coating was improved due to the higher harness of TiC particles in the matrix.

G. Xu et al. (2006) [14] used laser and TIG cladding techniques for surface modification of stellite-6 powder on the SUS403 substrate. The hardness value of the laser clad surface, TIG clad surface and substrate were 670, 420, 200 (Hv50) respectively. The specific wear rate of the laser clad surface, TIG clad surface and substrate were $0.8x10^{-8}$, $1.8 x10^{-8}$ and $2.3x10^{-8}$ mm^2/kg respectively. The highest hardness and lowest specific wear rate was obtained for the laser clad surface. Among the various cladding process, TIG arc surfacing is a cost-effective process. It produces high thickness surface composite layer which comprises of intermetallic compounds. Metallurgical bonding nature is the specific advantage of TIG arc surfacing [10]. The Recent studies on the TIG arc process is presented in Table 1.

3. SUMMARY AND FUTURE OUTLOOK

In this review article a comprehensive review on surface modification through TIG arc method is presented. The mechanical properties such as tensile strength, hardness & flexural strength is increased due to formation of intermetallic phases. The tribological, mechanical and corrosive properties of modified composite surface is improved due to uniform mixing of reinforcement particles with matrix. TIG arc method is a straightforward and movable method in comparison with conventional heat treatment methods hence, it is convenient method for alter the selective region.

ACKNOWLEDGEMENT

The author is thankful to Prof. Vipin, HOD & Professor, Department of Mechanical Engineering, Delhi Technological University, Delhi for their valuable guidance to execute the work.

REFERENCES

[1] A. Matthews, R.J. Hartley, P. Holiday, P.R. Stevenson, "The UK Engineering Coatings Industry in 2005," University of Hull, Hull, 1992.
[2] C.M. Cotell, J.A. Sprague, "Preface, Surface Engineering," ASM Handbook, ASM International, USA, 1994.
[3] R.F. Bunshah, "Handbook of Deposition Technologies for Films and Coatings: Science, Technology and Applications," 2nded. Noyes Publications, New York, 1994.
[4] J.P. Celis, D. Dreesa, M.Z. Huq, P.Q. Wu, M. DeBonte, "Hybrid processes—a versatile technique to match process requirements and coating needs. Surface and Coatings Technology," vol. 113, pp. 165-181, 1999.
[5] A.V. Muley, S. Aravindan, I.P. Singh, "Nano and hybrid aluminum based metal matrix composites: an overview," Manufacturing Review, vol. 15, pp. 1-13, 2015.
[6] J. Gandra, H. Krohn, R. M. Miranda, P.Vilaca, L. Quintino, J.F. dos Santos, Friction surfacing- A review, Journal of Materials Processing Technology, vol. 214 pp. 1062–1093, 2014.

[7] S.H. Choo, S. Lee, S.J. Kwon, "Effect of Flux Addition on the Microstructure and Hardness of TiC-Reinforced Ferrous Surface Composites Fabricated by High-Energy Electron Beam Irradiation," Metallurgical and Materials Transactions A, vol. 30A, pp. 3131-3141, 1999.
[8] L. Dubourg, F. Hlawk, A. Cornet, "Study of aluminium - copper–iron alloys: application for laser cladding," Surface and Coatings Technology, vol. 151–152, pp. 329-332, 2002.
[9] S. Jiguo, D. Wei, T. Wenda, Z. Di R. Jialie, "Dilution rate and microstructure of TIG arc Ni-Al powder surfacing," Frontier Mechanical Engineering, vol. 2, pp. 20-24, 2007.
[10] S. Islak, S. Buytoz, M. Karagoz, "Microstructural development on AISI 1060 steel by FeW/B_4C composite coating produced by using tungsten inert gas (TIG) process," Indian Journal of Engineering and Materials Sciences, vol. 19, pp. 253-259, 2012.
[11] S. Mridha, H.S. Ong, L.S. Poh, P. Cheang, "Intermetallic coatings produced by TIG surface melting," Journal of Materials Processing Technology, vol. 113, pp. 516-520, 2001.
[12] O.N. Celik, M. Ulutan, H. Gasan, U. Er. S. Buytoz, "Effects of graphite content on the microstructure and wear properties of an AISI8620 steel surface modified by tungsten inert gas (TIG)," Surface and Coatings Technology, vol. 206, pp. 1423-1429, 2011.
[13] W. Xinhong, Z. Zengd, S. Sili, Q. Shiyao, "Microstructure and wear properties of insitu TiC/FeCrBSi composite coating prepared by gas tungsten arc welding," Wear, vol. 260, pp. 25-29, 2006.
[14] G. Xu, M. Kutsun, Z. Liu, K. Yamada, "Comparison between diode laser and TIG cladding of Co-based alloys on the SUS403 stainless steel," Surface and Coatings Technology, vol. 201, pp.1138-1144, 2006.
[15] A. Ardeshiri, M. Heydarzadeh Sohi, A. Safaei, "Surface alloying of A2618 aluminum with silicon and iron by TIG process," Surface & Coatings Technology, vol. 310, pp. 87–92, 2017.
[16] S. Kumar, P.K. Ghosh, R. Kumar, "Surface modification of AISI4340 steel by multi-pass TIG arcing process," Journal of Materials Processing Technology, vol. 249, pp. 394–406, 2017.
[17] R. Kumar, P.K. Ghosh, S. Kumar, "Thermal and metallurgical characteristics of surface modification of AISI 8620 steel produced by TIG arcing process," Journal of Materials Processing Technology, vol. 240, pp. 420–431, 2017.
[18] Sh. Zangeneha, H.R. Lashgari, H.F. Lopezc, H.K. Farahani, "Microstructural characterization of TIG surface treating in Co-Cr-Mo-C alloy," Materials Characterization, vol. 132, pp. 223–229, 2017.
[19] A. Amirsadeghi, M. Heydarzadeh Sohi, S.F. Kashani Bozorg, "Effects of TIG Surface Melting and Chromium Surface Alloying on Microstructure Hardness and Wear Resistance of ADI," Journal of Iron and steel Research, vol.15, pp. 86-94, 2008.
[20] P.K. Ghosh, R. Kumar, "Surface Modification of Micro-Alloyed High-Strength Low-Alloy Steel by Controlled TIG Arcing Process," Metallurgical and Materials Transactions A, vol. 46A, pp. 831-842, 2015.
[21] M. Tavoosi, S. Arjmand, "In situ formation of Al/Al3 Ti composite coating on pure Ti surface by TIG surface process," Surfaces and Interfaces, vol. 8, pp. 1–7, 2017.
[22] A. Monfared, A.H. Kokabi, S. Asgari, "Microstructural studies and wear assessments of Ti/TiC surface composite coatings on commercial pure Ti produced by titanium cored wires and TIG process," Materials Chemistry and Physics, vol. 137, pp. 959-966, 2013.
[23] M.B. Karamıs, K. Yildizli, "Surface modification of nodular cast iron: A comparative study on graphite elimination," Materials Science and Engineering A, vol. 527, pp. 5225–5229, 2010.
[24] F.T. Cheng, K.H. Lo, H.C. Man, "NiTi cladding on stainless steel by TIG surfacing process Part I. Cavitation erosion behaviour," Surface and Coatings Technology, vol. 172, pp. 308–315, 2003.
[25] P.H. Lailatul, M.A. Maleque, "Surface Modification of Duplex Stainless Steel with SiC Preplacement Using TIG Torch Cladding," Procedia Engineering, vol. 184, pp. 737 – 742, 2017.
[26] N. Yuvaraj, S. Aravindan, Vipin, "Comparison studies on mechanical and wear behaviour of fabricated aluminium surface nano composites by fusion and solid-state processing," Surface & Coatings Technology, vol. 309, pp. 309-319, 2017.

Table 1. Summary of the investigations on TIG arc surfacing

Investigator Name	**Material investigated**	**Characteristic studied**	**Prominent Results**
A. Ardeshiri, et al., (2017) [15]	A2618- matrix material 30wt%Fe+70wt%Si	Microhardness and wear rate of the surface composite.	Hardness and wear resistance of the surface was improved due to formation of hard intermetallic compounds such as $Al1_3Fe_4$, Fe5Si3, Al_8Fe_2Si, and Al_5FeSi on the surface.
S. Kumar et al., (2017) [16]	AISI 4340 steel Surface modification	Effect of TIG arc passes and arcing current on the surface properties.	Multi pass TIG arcing on the surface improves the hardness and flexural strength of the material. The optimum arcing current and arc travel speed generated the appropriate heat input in order to improve the surface properties of the material.
R. Kumar et al., (2017) [17]	AISI 8620 low alloy steel	Metallurgical characteristics of modified surface using different TIG arcing and heat input.	TIG arcing process is clean, versatile and economic process for improve the surface modification. The metallurgical transformations such as Martensite, bainite and proeutectoid ferrite took place on the HAZ and modified fusion region.
Sh. Zangeneha et al., (2017) [18]	Cast Co-28Cr-5Mo-0.3C alloy	Formation of nano scale layer on the surface and microhardness.	Formation of nano scale M23C6 carbides on the surface. The higher amount of a thermal martensite formed on the surface are responsible for improvement of the surface hardness.
A Amirsadegh et al., (2008) [19]	Austempered Ductile cast iron (ADI) & Chromium surface alloying on ADI	Microhardness and wear properties.	The formation of ledeburiitic structure on the surface enhances the hardness and wear resistance. Chromium alloying ADI surface further improved the harness and wear resistance.
P.K.Ghosh and R. Kumar, (2014) [20]	Micro-alloyed HSLA steel	Effect of autogenous Conventional (C-TIGA) and pulse current tungsten (P-TIGA) inert gas arcing process	More depth of surface modification occurred by use of P-TIGA process in compare with C-TIGA process. Multi run P-TIGA process increases the hardness with wider surface modification.

M. Tavoosi and S. Arjmand, (2017) [21]	Pure Titanium with Al 1100 grade cladding	Al/Al3Ti composite coating on Ti surface. Wear resistance and hardness of the cladded surface.	The hardness of the cladded surface was improved due to formation of Al3Ti intermetallic compound on the surface. The hardness and corrosive resistance of the composite was increasing with increase in annealing time.
A. Monfared et al., (2013) [22]	Pure titanium & TiC particles for composite coating	Hardness and wear rate of the composite surface	Uniform microstructure was formed through the depth of coatings. The hardness (1100Hv) of the composite surface was improved 7 times more than substrate. Homogeneous dispersion of hard TiC particles in the ductile matrix enhances the wear resistance of the Ti/TiC composite surface.
M.B. Karamıs and K. Yıldızlı (2010) [23]	Nodular Cat Iron (NCI)	Surface modification through elimination of graphite nodules on the surface.	The surface was modified with required depth by controlling the process parameters in order to modify the nodular graphite structure.
F.T. Cheng et al., (2003) [24]	AISI316 stainless steel & NiTi wire for cladding	Dilution rate and hardness of the cladded surface.	Thick NiTi laye with Fe and Cr was formed on the surface through strong interfacial bonding between the NiTi and steel substrate. The hardness (700Hv) of the cladded surface was higher than base Stainless-steel hardness (200Hv).
S. Mridha et al., (2001) [11]	Commercial pure Titanium(Cp-Ti)	Surface hardness	The surface hardness(2000Hv) of the surface modified surface was improved due to formation of titanium nitrides on the surface.
P.H. Lailatul and M.A. Maleque, (2017) [25]	Duplex Stainless Steel with SiC powder for coating	Effect of welding process parameters on hardness of modified surface.	The hardness of the modified surface was improved 2 to 3 times than base alloy. The dendrite microstructure was formed during melting of steel and uniform mixing of SiC particles with the steel are the main reasons for enchantment of hardness.
N.Yuvaraj et al., (2017) [26]	Al 5083 with B_4C nano particles	Hardness and wear properties.	The hardness and wear properties of the surface composite fabricated through TIG arc process was improved due to uniform dispersion of reinforcement particles in the matrix.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

Synthesis and Characterization of Thiosemicarbazone Metal Complexes

Shankar Suman
Department of Applied Chemistry, Delhi Technological University, Delhi, India
Ram Singh*
Department of Applied Chemistry, Delhi Technological University, Delhi, India

ABSTRACT

The coordination chemistry of thiosemicarbazonesplays an important role in a wide range of biological, pharmaceutical and material activities. In many cases, the activities of thiosemicarbazones increases many folds after their metalation. This paper discusses the synthesis and characterization of thiosemicarbazone metal complexes having focus on transition metals.

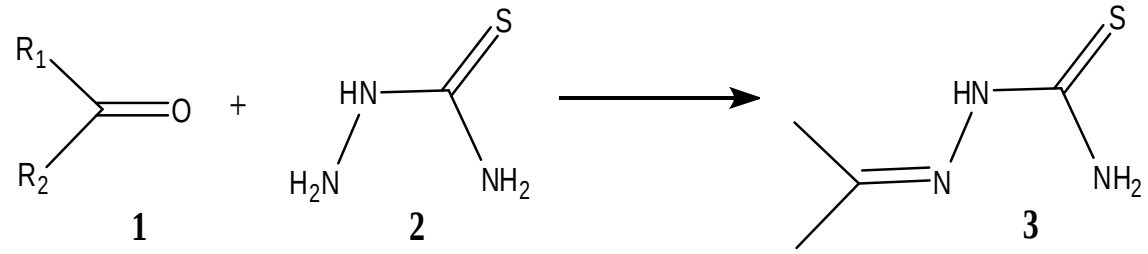
KEYWORDS

Semicarbazones, thiosemicarbazones, metal complexes,* Schiff base, *ketimines, aldimines

1. INTRODUCTION

Coordination compounds have shown their importance in varied walks of life since their discovery due to their wide applicability.[1]The electronegative atoms like nitrogen, sulphur or oxygen plays important role in the formation of coordination compounds. Out of many versatile ligands possessing these atoms, thiosemicarbazones(**3**) are very adaptable and multipurpose ligands.The preparation of thiosemicarbazones (**3**) takes place by reacting the thiosemicarbazides (**2**) and suitable aldehydes or ketones (**1**) (Scheme 1).[2,3]For aldehydes, the hydrogen atom is one of the substituents in the structure **1**, while second substituent may be an alkyl, an aryl or a heterocyclic group. For the ketones, both the substituents may be an alkyl, an aryl or a heterocyclic group.

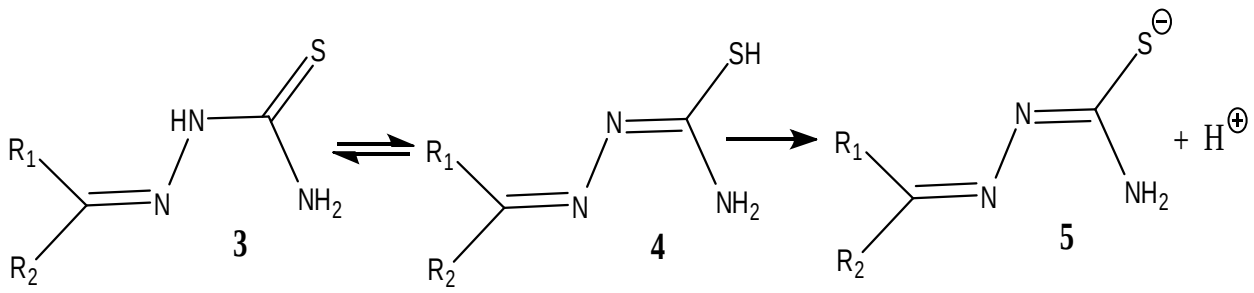
Thethiosemicarbazones (**3**)have been utilized for complex formation due to high coordination tendency and better selectivity.[2]When this ligand coordinate with metal, it either interact as neutral molecule (**3**) or as anionic ligands(**5**) (Scheme 2). The bonding usually takes place by using hydrazine nitrogen atom and the sulphur atom.[3]The metal complexes of **3** have displayed their utility in different types of industries like pharmaceutical, agricultural, chemical, polymers and so on.[4-9] The metal complex formation of **3** have been initiated in the early sixties only and later a number of complexes synthesized and studied.[9,10] This paper deals with synthesis and characterization of transition metal complexes of thiosemicarbazones.



1 2 3

Scheme 1. Synthesis of thiosemicarbazones (**3**)

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018



Scheme 2. Tautomerization and anion formation

2. SYNTHESIS AND CHARACTERIZATION OF THIOSEMICARBAZONE METAL COMPLEXES

The synthesis of the ligand **3** have been achieved by traditional as well as environmental methods.[11-14]One of the most general method is to dissolve the thiosemicarbazide in methanol (CH_3OH) or aqueous methanol followed by addition of a given aldehyde or ketone. The reaction mixture is refluxed for 10 to 12 hours. The completion of the reaction is checked with thin layer chromatography (TLC). The evaporation of the solvent gives crude thiosemicarbazone which further recrystallized with methanol to get the pure product. The reaction was also performed by direct refluxing[11] or using microwave irradiation[12].

The formation of metal complexes of **3** depends upon the type of metal employed. The metals have been used as their salts like metal halides, metal nitrates, metal acetates, metal sulfates etc.[9] The ligand and the metal complex have been taken in appropriate solvent and further either refluxed or stirred at appropriate temperature to form the complexes.[15-18] Its difficult to generalize a method for complexation of all types of metals. There were many metals that requires different starting materials like vanadium.[19] The formation of Cr(III) complex was reported from $CrCl3\cdot 6H_2O$ in aqueous medium.[20]Some examples of the formation of metal complexes are given below followed by the methods of characterization.

Khan et al[13] synthesized metal complexes of thiosemicarbazone derivative using the salts of copper, nickel and cobalt as their chloride or acetate where coordination was reported through the azomethine nitrogen and thionic Sulphur (Figure 1a).Shawish et al[21]have synthesized the Ni(II) complexes of dihydroxybenzaldehydethiosemicarbazones and 1,2-bis(diphenylphosphino)ethane. The complexes possess have a Square-planar geometry. It was observed that the ligand acted as tridentate ligand with coordination through S, O, and N atoms. In another study, the Ni(II) complexes of 4-(*p*-substituted phenyl)thiosemicarbazonesof salicylaldehyde have been reported (Figure 1b& c).[22]The study observed that the redox potential of the complexes have been affected by the substitution at 4-postion in the phenyl ring. The metal complexes of Fe(III), Ni(II), Hg(II) and Zn(II) with ligand benzyl α-monoxime pyridyl thiosemicarbazonehave been synthesized by El-Reash et al.[23]The salts of the metal and the ligand were refluxed in absolute ethanol to get the complexes that were found to be stable in air. The chloride of Fe(III), Ni(II), Hg(II) and acetate of Zn(II) were used for the reaction.[23]

The copper(II) complexes of isopropylmethyl ketone thiosemicarbazone and 4-aminoacetophenonethiosemicarbazone have been synthesized by refluxing the ligand and the hydrated metal salt in aqueous ethanolic solution.[24]The metal complexes of Co(II), Ni(II) and Cu(II) with thiosemicarbazone (L_1) derived from pyrole-2-carboxyaldehyde have been prepared by refluxing the reaction mixture in ethanol.[25]The synthesized metal complexes $Co(L_1)_2(NO_3)_2$ and $Ni(L_1)_2(NO_3)_2$ showed tetrahedral geometry whereas Ni(L_1)2(NO3)2 possess tetrahedral geometry. The metal complexes are of high-spin type.[25]

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

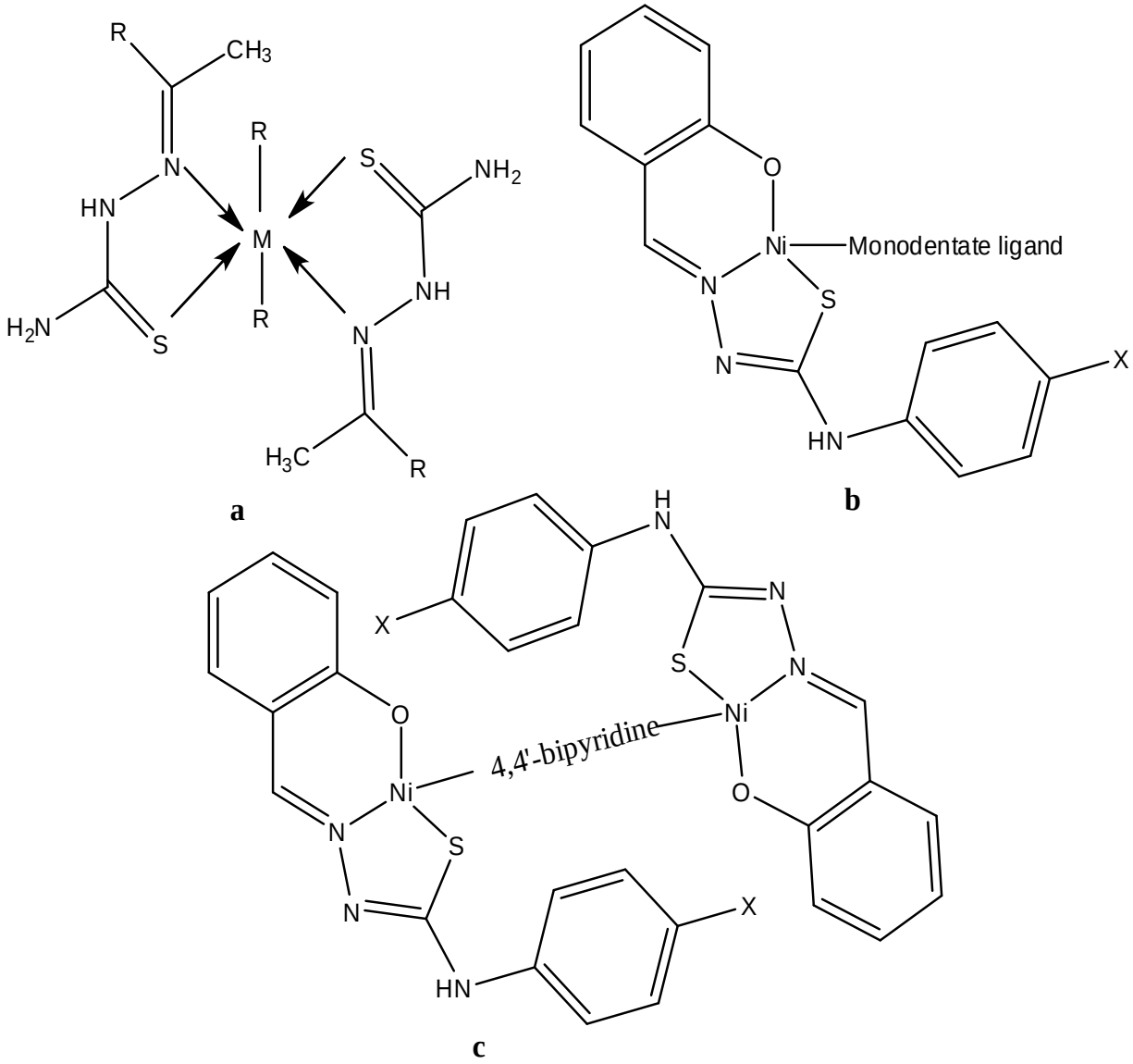


Figure 1. General Structures of metal complexes

The preparation of metal ccomplexes belonging to 4d and 5d transition metal series were done by different routes depending upon the type of metal salt utilized.[26-28]For the preparation of molybdenum (Mo) complex, $MoO_2(acac)_2$ was used and then the ligand thiosemicarbazone was replaced with acac.[26] For metals like Tungsten (W) and Rhenium (Re), the metal carbonyls such as $W(CO)_6$[27] and $ReX(CO)_3(CH_3CN)_2$[28] were used with the ligand.

The characterization of metal complexes of thiosemicarbazones have usually been done with the help of UV-visible, FTIR, NMR, and ESR spectroscopic methods and elemental analysis.[9] In FTIR, the ligand and the corresponding metal complex exhibit characteristic bands depending upon the bonding behaviour. The frequency for the presence of N^3H group (Figure 2) in FTIR suggests the coordination of neutral ligand (**3**) to the metal ion. Its absence, shows the involvement of ligands in the form of **4** or **5** (Scheme 2).The anionic ligand complexes exhibit vibrational modes due to the N^1H group (Figure 2) only, in the rangeof 3300–3500cm^{-1}.[9]

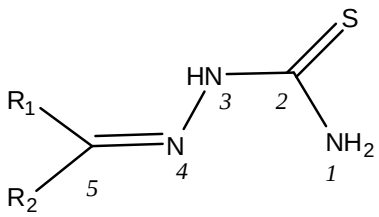


Figure 2. The ligand with atom marked

3. CONCLUSIONS

Thiosemicarbazonesare versatile ligands and have always been a focus of research due to their variable donor properties and diversity in structure with respect to the carbonyl compounds utilized. The bonding for the formation of complexesusually takes place through the sulphur and hydrazine nitrogen atoms. The thiosemicarbazone metal complexes have shown potential biological activities and ion sensing properties. Their synthesis usually performed by refluxing the ligand and metal salts in ethanol or aqueous ethanol as solvent of choice. The denticity of the ligands depends upon the substituents present with thiosemicarbazone. The formation of the metal complexes has been characterized with UV-visible, FTIR, NMR and elemental analysis.

4. REFERENCES

[1] Lawrance, G.A. 2010. Introduction to Coordination Chemistry, John Wiley & Sons. ISBN 9780470519301.
[2] Shalin, K., Dhar, N.D. and Sharma N.P. 2009. Application of metal complexes of Schiff, bases, J. Sci. Indust. Res., 68, 181-187.
[3] Pelosi, G. 2010. Thiosemicarbazone metal complexes: From structure to activity, The Open Crystal. J., 3, 16-28.
[4] Farrell, N. 2002.Biomedical uses and applications of inorganic chemistry. An overview. Coord. Chem. Rev. 232, 1-4.
[5] Zahra, A.,Ekk, S., Weisheng, L., Yinfa, M., Charles, C. and Subhash, P. 2005. Nickel (II) complexes of naphthaquinonethiosemicarbazone and semicarbazone: Synthesis, structure, spectroscopy and biological activity. J. Inorg. Biochem. 99, 1526-1531.
[6] West, D.X., Padhye, S.B. andSonawane, P.B. 1991. Structure and Bonding, Springer-Verlag: New York.
[7] Atalay, T. andAkgemci, E.G.1998. Thermodynamic studies of some complexes of 2-benzoylpyridine 4-phenyl-3-thiosemicarbazone.Tr. J. Chem. 22, 123-127.
[8] Kolb, V.M.,Stupar,J.W.,Janota,T.E. andDuax,W.L. 1989. Abnormally high IR frequencies for the carbonyl group of semicarbazones of the benzaldehyde and acetophenone series. J. Org. Chem. 54, 2341-2346.
[9] Lobana, T.S., Sharma, R.,Bawa, G. and Khanna, S. 2009. Bonding and structure trends of thiosemicarbazone derivatives of metals - An overview. Coord. Chem. Rev. 253, 977–1055.
[10] Gingras, B.A.,Somorjai, R.L. and Bayley, C.H. 1961.The preparation of some thiosemicarbazones and their copper complexes. Can. J. Chem. 39, 973-985.
[11] Chandra, S.,Raizada, S., Tyagi, M. and Sharma, P.K. 2008.Spectroscopic and biological approach of Ni(II) and Cu(II) complexes of 2-pyridinecarboxaldehyde thiosemicarbazone. Spectrochim. Acta Part A, 69, 816-821.
[12] Tadjarodi, A. and Najjari, S. 2012.Preparation of 2-pyridinecarbaldehyde thiosemicarbazone by microwave irradiation. International Electronic Conference on synthetic organic chemistry, c002.
[13] Khan, S.A.,Asiri, A.M., Al-Amry, K. and Malik, M.A. 2014. Synthesis, characterization, electrochemical studies, and *In-vitro* antibacterial activity of novel thiosemicarbazone and its Cu(II), Ni(II), and Co(II) complexes. The Scientific World J. 2014, Article ID 592375, 9 pages.
[14] Scovill, J.P. 1991.A facile synthesis of thiosemicarbazidesand thiosemicarbazones by thetransamination of 4-methyl-4-phenyl-3-thiosemicarbazide. Phosphorus, Sulfur, Silicon 60, 15-19.
[15] West,D.X.,Liberta,A.E.,Padhye,S.B.,Chikate,R.C.,Sonawane,P.B.,Kumbhar, A.S. and Yerande, R.G. 1993. Thiosemicarbazonecomplexes of copper(II): structural and biological studies. Coord. Chem. Rev. 123, 49–57.
[16] Duan,C.-Y., Liu,Z.-H., Shi,Y.-C. and You,X.-Z. 1999.Synthesis, crystal structure and nonlinear optical properties of thiosemicarbazone zinc complex. J. Coord. Chem. 47, 433-439.

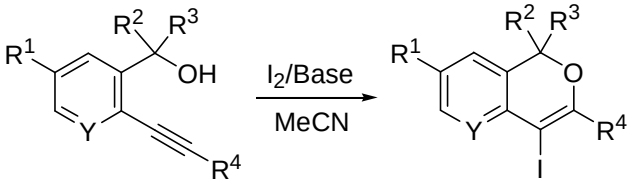
[17] Casas,J.S., Garcia-Tasende,M.S. andSordo, J. 2000.Main group metal complexes of semicarbazones and thiosemicarbazones. A structural review. Coord. Chem. Rev. 209 (2000) 197-261.
[18] Casas,J.S., Garcia-Tasende,M.S. andSordo, J. 1999. Structural aspects of the coordination chemistry of organothallium(III) and organomercury(II) derivativesCoord. Chem. Rev. 193-195, 283-359.
[19] Milanesio,M.,Viterbo, D., Hernandez,R.P., Rodriguez,J.D., Ramirez-Ortiz, J. and Valdes-Martinez, J. 2000.Synthesis, characterization and novel crystal structure of (salicylaldehyde-4-phenylthiosemicarbazidato) ammonium dioxovanadate(V) with a V-S bond. Inorg. Chim. Acta 306, 125-129.
[20] Sonawane, P.,Chikate, R.,Kumbhar, A.,Padhye, S. andDoedens, R.J. 1994.Inequivalent coordination of thiosemicarbazone ligands in cobalt (III) and chromium (III) complexes. Polyhedron 13, 395-401.
[21] Shawish, H.B.,Maah, M., Halim, S.N.A. and Shaker, S.A. 2016. Synthesis, characterization and structural studiesof binuclear nickel(II) complexes derivedfrom dihydroxybenzaldehydethiosemicarbazones,bridged by 1,2-bis(diphenylphosphino)ethane.Arab. J. Chem. 9, S1935–S1942.
[22] Saswati, Dinda, R.,Schmiesing, C.S., Sinn, E.,Patil, Y.P.,Nethaji, M., Stoeckli-Evans, H. andAcharyy, R. 2013. Mixed-ligand nickel(II) thiosemicarbazone complexes: Synthesis, characterizationand biological evaluation. Polyhedron 50, 354–363.
[23] Bedier, R.A., Yousef, T.A., Abu El-Reash, G.M. and El-Gammal, O.A. 2017. Synthesis, structural, optical band gap and biological studies on iron(III), nickel (II), zinc (II) and mercury (II) complexes of benzyl amonoximepyridyl thiosemicarbazone. J. Mol. Str. 1139, 436-446.
[24] Chandra, S. and Gupta, L.K. EPR, mass, IR, electronic, and magnetic studies on copper(II) complexes of semicarbazones and thiosemicarbazones, Spectrochimica Acta Part A 61 (2005) 269–275.
[25] Chandra, S. and Kumar, A. 2007. Electronic, epr and magnetic studies of Co(II), Ni(II) and Cu(II) complexes with thiosemicarbazone (L1) and semicarbazone (L2) derived from pyrole-2-carboxyaldehyde.Spectrochim.Acta Part A 67, 697–701.
[26] Stelzig, L., Kotte, S. and Krebs, B. 1998.Molybdenum complexes with tridentate NS_2 ligands. Synthesis, crystal structures and spectroscopic properties. J. Chem. Soc. Dalton Trans. 2921-2926.
[27] Valdes-Martinez, J., Enriquez, A., Cabrera, A. and Espinosa-Perez, G. 1996. CH…O hydrogen-bonding interactions as packing tools. Crystal structure of [$W(CO)_5$(acetone-4-methylthiosemicarbazone)]. Polyhedron 15, 897-901.
[28] Carballo, R., Casas, J.S., Garcia-Martinez, E.,Pereiras-Gabian, G., Sanchez, A., Sordo, J. and Vazquez-Lopez,E.M. 2003. Rhenium(I)-induced cyclization of thiosemicarbazonesderived from β-Keto Esters. Inorg. Chem. 42, 6395-6403.

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

Synthetic Strategies for Isochromenes: Recent Advances

Saurabh Mehta*
Department of Applied Chemistry, Delhi Technological University,
Bawana Road, Delhi, India

ABSTRACT



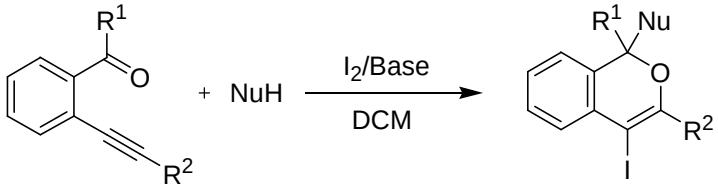
Y = CH, N; R^1 = H, OMe, NO_2; R^2, R^3 = H, alkyl; R^4 = alkyl, aryl, thienyl

A brief review of the recent synthetic methodologies for Isochromenes has been presented. Various functionalized substrates such as substituted alkynes are conveniently cyclized to afford diverse isochromenes with interesting substitution patterns. The use of external nucleophiles in some cases adds additional chemical diversity. Finally the metal-catalyzed dimerization and trimerization strategies lead to the formation of more complex and skeletally diverse spiro-isochromenes.

INTRODUCTION

Isochromenes are important heterocyclic compounds, and there are many examples of natural products and biologically active compounds known in the literature that contain this ring system.[1,2] Many synthetic research groups have previously reported interesting synthetic strategies for isochromenes through the annulation of acyclic starting materials.[3,4] The Electrophile mediated cyclization, particularly iodocyclization of functionally substituted alkynes has been extensively explored for the synthesis of a variety of heteroaromatic compounds.[5,6] These include important heterocyclic scaffolds such as indoles,[7,8] benzofurans,[9] chromones,[10] furanones,[11] benzothiophenes,[7,12] isoxazoles,[13] isoindolinones,[14] quinolines,[15,16] *etc*. Larock et al. and Barluenga group have previously accomplished the synthesis of iodoisochromene derivatives through the external nucleophile-mediated iodocyclization of 2-alkynyl benzaldehyde substrates (Scheme 1).[17,18] Here we describe the recent advancements made in the synthesis of this important heterocyclic scaffold.

Scheme 1. Nucleophile-mediated Synthesis of Isochromenes.



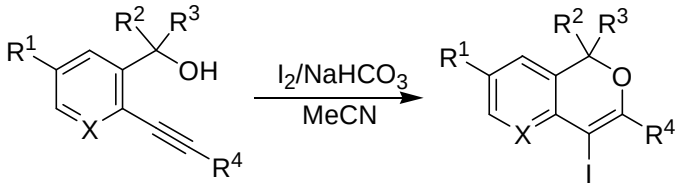
R^1 = H, Me; R^3 = alkyl, aryl, heteroaryl
NuH = MeOH, EtOH, $PhNMe_2$, PhOH, *etc.*

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

DISCUSSION

We have previously reported an efficient methodology for the direct synthesis of 3-iodo-isochromenes, through the electrophilic cyclization of substituted 2-(1-alkynyl)benzylic alcohols in very mild reaction conditions (Scheme 2).[19]

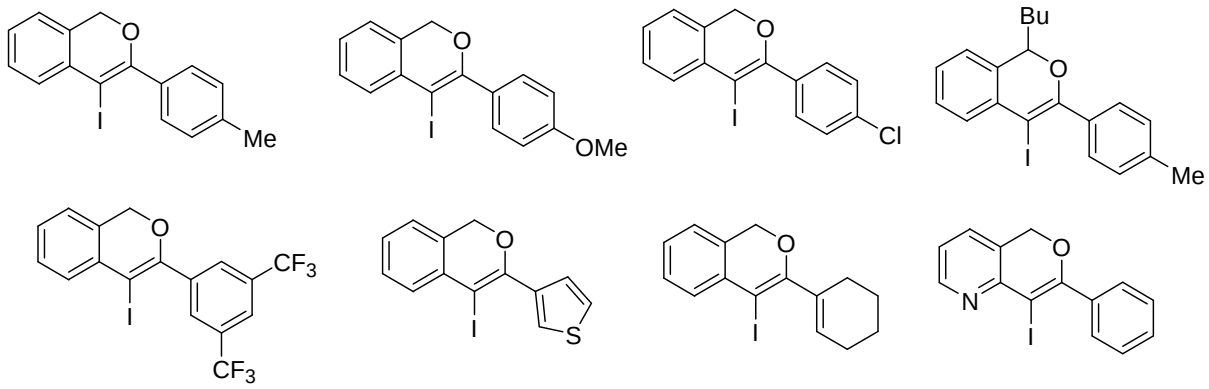
Scheme 2. Synthesis of Isochromenes through the electrophilic cyclization of substituted 2-(1-alkynyl)benzylic alcohols.



X = CH, N; R^1 = H, OMe, NO_2; R^2, R^3 = H, alkyl; R^4 = alkyl, aryl, thienyl

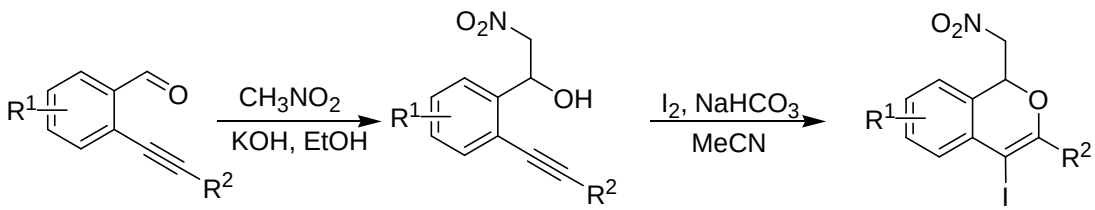
Many iodine-containing derivatives of isochromenes were synthesized using this methodology. The cyclized products were characterized by IR, MS, ^{1}H NMR, and ^{13}C NMR spectroscopy. Figure 1 below lists a few representative Isochromene derivatives.

Figure 1: Representative Isochromenes synthesized through Iodocyclization strategy.



Kundu and coworkers reported another interesting synthesis of Isochromenes involving the electrophilic cyclization (Scheme 3).[20] The alkynyl aldehydes were first reacted with nitromethane and the resulting alcohols were iodocyclized to give the Isochromene derivatives in good to moderate yields.

Scheme 3. Synthesis of Isochromenes from 2-Alkynyl benzaldehydes.

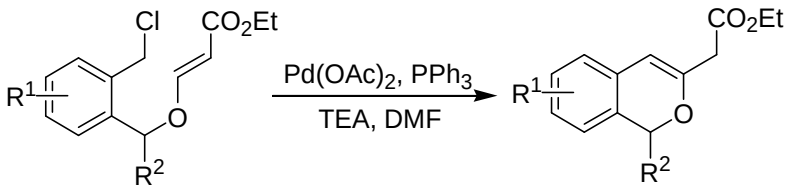


Gharpure et al. reported an intramolecular Heck reaction involving a benzylic C(sp^3)-C(sp^2) bond formation in β-alkoxyacrylates for the synthesis of isochromenes (Scheme 4).[21] Competitive Heck reaction

International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

between normal olefin and vinylogous carbonate moiety led to dihydronaphthalene product via coupling with olefin exclusively. The methodology was also extended for the synthesis of the core of monocerolides.

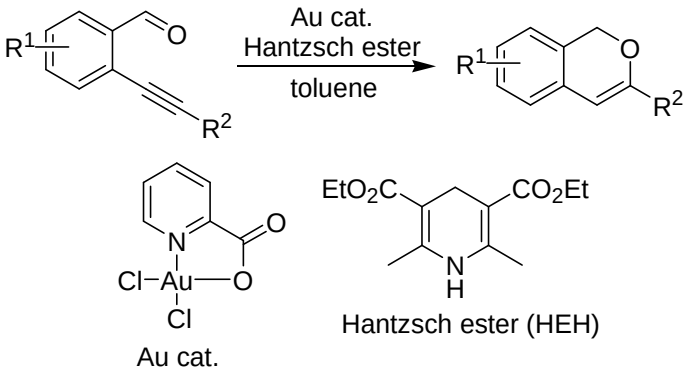
Scheme 4. Synthesis of Isochromenes from therough Intramolecular Heck coupling.



R^1 = H, Functional Groups; R^2 = alkyl, aryl, *etc.*

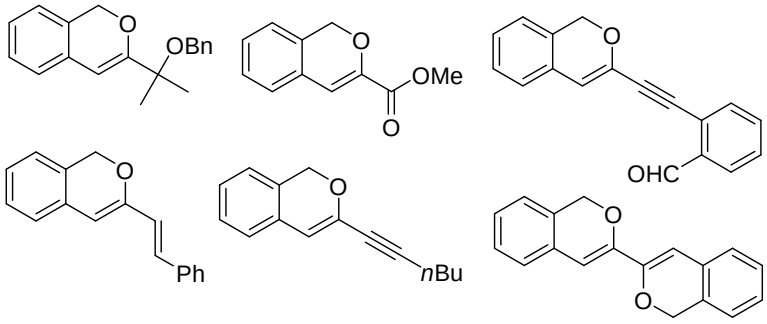
Michelet et al. reported an interesting Synthesis of 1*H*-Isochromenes through a Gold-catalyzed Domino Cycloisomerization/Reduction strategy (Scheme 5).[22] The authors demonstrated that the commercially available catalyst complex associated with the Hantzsch ester (HEH) allowed, easy and efficient access to diverse isochromenes.

Scheme 5. Gold-catalyzed Domino Cycloisomerization/Reduction strategy.



The same research group later established the possible mechanism for the reaction and also expanded the scope of the methodology, and successfully demonstrated that the important functional groups such as aldehyde, ester, alkene, alkyne, etc. were tolerated. (Figure 2).[23]

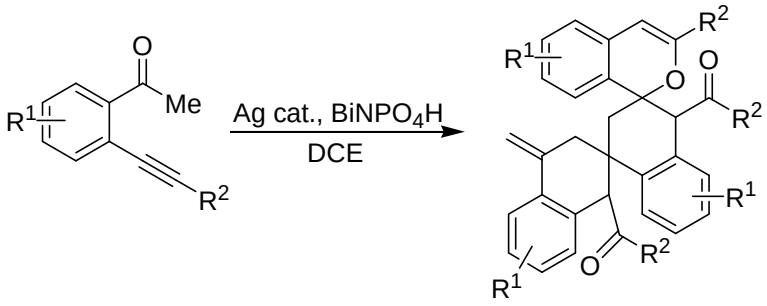
Figure 2. Scope of the Cycloisomerization/Reduction strategy.



International Journal of Engineering Technology Science and Research
IJETSR
www.ijetsr.com
ISSN 2394 – 3386
Volume 5, Issue 4
April 2018

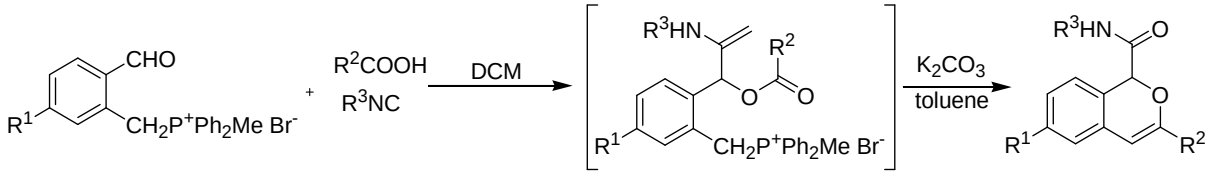
Jiang et al. reported an interesting silver-catalyzed tunable dimerization and trimerization strategy involving β-alkynyl ketones, for the synthesis of skeletally diverse spiro-isochromenes (Scheme 6).[24] The single crystal X-ray analyses of the products confirmed the structures of spiro isochromene derivatives.

Scheme 6. Synthesis of Spiro-isochromenes.



Ding et al. recently reported a multicomponent reaction for the synthesis of diverse Isochromenes (Scheme 7).[25] Thus the reaction of a phosphonium salt, acid and isocyanide afforded the corresponding 1*H*-isochromenes in good yields.

Scheme 7. Multicomponent Reaction.



CONCLUSIONS

Thus, a variety of efficient ways for making Isochromene scaffold have been reviewed. These include metal-catalyzed, halogen-mediated annulations as well as multicomponent synthetic routes. Such strategies are of great importance and interest to the synthetic as well as medicinal chemists.

REFERENCES

(1) Kanokmedhakul, S.; Kanokmedhakul, K.; Nasomjai, P.; Louangsysouphanh, S.; Soytong, K.; Isobe, M.; Kongsaeree, P.; Prabpai, S.; SuSomdej, A.; Kwanjai, K.; et al. Antifungal Azaphilones from the Fungus Chaetomium Cupreum CC3003. *J. Nat. Prod.* **2006**, *69*, 891–895.
(2) Gao, J.-M.; Yang, S.-X.; Qin, J.-C. Azaphilones: Chemistry and Biology. *Chem. Rev.* **2013**, *113*, 4755–4811.
(3) Butin, A. V.; Abaev, V. T.; Mel'chin, V. V.; Dmitriev, A. S.; Pilipenko, A. S.; Shashkov, A. S. A Furan Recyclization Reaction as a New Approach to Isochromenes. *Synthesis* **2008**, *11* (11), 1798–1804.
(4) Han, Y.; Wang, X.; Wang, X.; Lv, L.; Diao, G.; Yuan, Y. Ligand-Free Copper Powder Catalyzed Direct Coupling Reaction of Heterocyclic C – H Bonds and Aryl Halides. *Synthesis* **2012**, *44*, 3–8.
(5) Godoi, B.; Schumacher, R. F.; Zeni, G. Synthesis of Heterocycles via Electrophilic Cyclization of Alkynes Containing Heteroatom. *Chemical Reviews*. American Chemical Society 2011, pp 2937–2980.
(6) Gabriele, B.; Mancuso, R.; Larock, R. C. Recent Advances in the Synthesis of Iodoheterocycles via Iodocyclization of Functionalized Alkynes. *Current Organic Chemistry* **2014**, *18* (3), 341–358.
(7) Danilkina, N. A.; Kulyashova, A. E.; Khlebnikov, A. F.; Bräse, S.; Balova, I. A. Electrophilic Cyclization of Aryldiacetylenes in the Synthesis of Functionalized Enediynes Fused to a Heterocyclic Core. *J. Org. Chem.* **2014**, *79* (19), 9018–9045.
(8) Li, Y. L.; Li, J.; Yu, S. N.; Wang, J. B.; Yu, Y. M.; Deng, J. A Concise Approach for the Synthesis of 3-Iodoindoles

and 3-Iodobenzo[b]furans via Ph<inf>3</inf>P-Catalyzed Iodocyclization. *Tetrahedron* **2015**, *71* (43), 8271–8277.

(9) Mehta, S. Benzofuran Synthesis through Iodocyclization Reactions: Recent Advances. *MOJ Biorg. Org. Chem.* **2017**, *1*, 45.

(10) Zhou, C.; Dubrovsky, A. V.; Larock, R. C. Diversity-Oriented Synthesis of 3-Iodochromones and Heteroatom Analogues via ICl-Induced Cyclization. *J. Org. Chem.* **2006**, *71*, 1626–1632.

(11) Just, Z. W.; Larock, R. C. Synthesis of 2(3H)-Furanones via Electrophilic Cyclization. *J. Org. Chem.* **2008**, *73*, 2662–2667.

(12) Aurelio, L.; Volpe, R.; Halim, R.; Scammells, P. J.; Flynn, B. L. Synthesis of Thieno-Fused Heterocycles through Reiterative Iodocyclization. *Advanced Synthesis and Catalysis* **2014**, *356* (9), 1974–1978.

(13) Okitsu, T.; Sato, K.; Potewar, T. M.; Wada, A. Iodocyclization of Hydroxylamine Derivatives Based on the Control of Oxidative Aromatization Leading to 2,5-Dihydroisoxazoles and Isoxazoles. *Journal of Organic Chemistry* **2011**, *76* (9), 3438–3449.

(14) Brahmchari, D.; Verma, A. K.; Mehta, S. Regio- and Stereoselective Synthesis of Isoindolin-1-Ones through BuLi-Mediated Iodoaminocyclization of 2-(1-Alkynyl)benzamides. *J. Org. Chem.* **2018**, *83*, 3339−3347.

(15) Ali, S.; Zhu, H.-T.; Xia, X.-F.; Ji, K.-G.; Yang, Y.-F.; Song, X.-R.; Liang, Y.-M. Electrophile-Driven Regioselective Synthesis of Functionalized Quinolines. *Org. Lett.* **2011**, *13*, 2598–2601.

(16) Sasaki, T.; Moriyama, K.; Togo, H. Preparation of 3-Iodoquinolines from N-Tosyl-2-Propynylamines with Diaryliodonium Triflate and N-Iodosuccinimide. *J. Org. Chem.* **2017**, *82*, 11727–11734.

(17) Barluenga, J.; Vázquez-Villa, H.; Ballesteros, A.; González, J. M. Cyclization of Carbonyl Groups onto Alkynes upon Reaction with IPy2BF4and Their Trapping with Nucleophiles: A Versatile Trigger for Assembling Oxygen Heterocycles. *Journal of the American Chemical Society* **2003**, *125* (30), 9028–9029.

(18) Yue, D.; Della Cá, N.; Larock, R. C. Syntheses of Isochromenes and Naphthalenes by Electrophilic Cyclization of Acetylenic Arenecarboxaldehydes. *Journal of Organic Chemistry* **2006**, *71* (9), 3381–3388.

(19) Mancuso, R.; Mehta, S.; Gabriele, B.; Salerno, G.; Jenks, W. S.; Larock, R. C. A Simple and Mild Synthesis of 1H-Isochromenes and (Z)-1-Alkylidene-1,3-Dihydroisobenzofurans by the Iodocyclization of 2-(1-Alkynyl)benzylic Alcohols. *J. Org. Chem.* **2010**, *75* (3), 897–901.

(20) Arigela, R. K.; Samala, S.; Mahar, R.; Shukla, S. K.; Kundu, B. Synthesis of Triazolo Isoquinolines and Isochromenes from 2-Alkynylbenzaldehyde via Domino Reactions under Transition-Metal-Free Conditions. *J. Org. Chem.* **2013**, *78*, 10476–10484.

(21) Gharpure, S. J.; Shelkea, Y. G.; Reddy, S. R. B. Synthesis of Isochromene Derivatives Using an Intramolecular Benzylic C(sp^3)–C(sp^2) Bond Forming Heck Reaction on Vinylogous Carbonates. *RSC Adv.* **2014**, *4*, 46962–46965.

(22) Tomás-Mendivil, E.; Starck, J.; Ortuno, J.-C.; Michelet, V. Synthesis of Functionalized 1H-Isochromene Derivatives via a Au-Catalyzed Domino Cycloisomerization/Reduction Approach. *Org. Lett.* **2015**, *17*, 6126–6129.

(23) Tomás-Mendivil, E.; Heinrich, C. F.; Ortuno, J.-C.; Starck, J.; Michelet, V. Gold-Catalyzed Access to 1H-Isochromenes: Reaction Development and Mechanistic Insight. *ACS Catal.* **2017**, *7*, 380–387.

(24) Wang, D.; Liu, S.; Lan, X.; Paniagua, A.; Hao, W.; Li, G.; Tu, S.; Jiang, B. Tunable Dimerization and Trimerization of β-Alkynyl Ketones via Silver Catalysis for Accessing Spiro and Dispiro Compounds Containing 1H-Isochromene. *Adv. Synth. Catal.* **2017**, *359*, 3186 –3193.

(25) Wang, L.; Guan, Z.-R.; Ding, M.-W. One-Pot Synthesis of 1H-Isochromenes and 1,2-Dihydroisoquinolines by a Sequential Isocyanide-Based multicomponent/Wittig Reaction. *Org. Biomol. Chem.* **2016**, *14*, 2413–2420.

Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning

Dhruv Rathi*, Sushant Jain**, Dr. S. Indu***

* Department of Computer Engineering, Delhi Technological University, New Delhi, India
Phone: +919868904222, E-Mail: dhruvrathi15@gmail.com, http://dhruvrathi.me/
** Department of Computer Engineering, Delhi Technological University, New Delhi, India
E-Mail: sjsushantjain22@gmail.com
*** Department of Electronics and Communication, Delhi Technological University, New Delhi, India E-Mail: s.indu@dce.ac.in, http://www.dtu.ac.in/Web/Departments/Electronics/faculty/sindu.php

***Abstract*— The target of this paper is to recommend a way for Automated classification of Fish species. A high accuracy fish classification is required for greater understanding of fish behavior in Ichthyology and by marine biologists. Maintaining a ledger of the number of fishes per species and marking the endangered species in large and small water bodies is required by concerned institutions. Majority of available methods focus on classification of fishes outside of water because underwater classification poses challenges such as background noises, distortion of images, the presence of other water bodies in images, image quality and occlusion. This method uses a novel technique based on Convolutional Neural Networks, Deep Learning and Image Processing to achieve an accuracy of 96.29%. This method ensures considerably discrimination accuracy improvements than the previously proposed methods.**

Keywords- Fish Species Classification, Deep Learning, Convolutional Neural Network, Morphological Operations, Otsu binarization, Otsu Thresholding, Pyramid Mean Shifting, Computer Vision

I. INTRODUCTION

Monitoring the behavior of different species of fishes is of primary importance for getting insights into a marine ecological system. The count and distribution of the various species of fishes can give valuable insights about the health of the ecological system and can be used as a parameter for monitoring environmental changes. Visual classifying of fishes can also help trace their movement and give patterns and trends in their activities providing a deeper [1]knowledge about the species as a whole.

The study of the behavior of the fishes can be automated by getting visual feedback from multiple locations and automating the process of visual classification of fishes, which will give significantly larger amounts of data for pattern recognition. Although, there have been many advances in classifying fish taken out of water [1][2][3][4] or in artificial conditions, such as in tanks with adequate lighting [5], there has been no significant breakthrough in the classification of fishes in datasets created from underwater videos. The challenges faced during underwater classification of fish species include noise, distortion, overlap, segmentation error and occlusion [6]. Also, the complex environment restricts simpler approaches like luminance, and background subtraction includes issues such as colors shifting, inconsistent lighting and presence of sediments in water and undulating underwater plants.

Fish species recognition is a multi-class classification problem and is a compelling research field of machine learning and computer vision. The state-of-the-art algorithms implemented over individual input images which perform classification mainly using shape and texture feature extraction and matching [7][8]. All the existing work either deals with a small dataset distinguishing between less number of species or has a low accuracy. Our proposed method uses Convolutional Neural Networks which makes the process simpler and more robust even while working with a large dataset. CNNs are also much more flexible and can adapt to the new incoming data as the dataset matures. We make use of the fish dataset from the Fish4Knowledge project [24] for testing our algorithm. We perform the classification by pre-processing the images using Gaussian Blurring, Morphological Operations, Otsu's Thresholding and Pyramid Mean Shifting, further feeding the enhanced images to a Convolutional Neural Network for classification.

The remaining paper is catalogued as follows: Section II. reviews the current research done in this field; Section III. delineates the proposed algorithm; Section IV. gives the experimental results; Section V. discusses the Future Works and Conclusion and Section VI. gives the References.

II. RELATED WORKS

There are numerous approaches proposed by researchers for the classification of fish species as delineated below:

C. Spampianto et al. [7] attempt to classify fishes using texture features extracted from the gabor filtering, gray level histogram and the histogram of Fourier descriptors of boundaries and Curvature Scale Space Transform used for shape features extracted using. The algorithm was tested on a dataset of 360 images and achieved an accuracy of 92%.

Takakazu Ishimatsu et al. [8] used two identification features: speckle patterns and scale forms of fish and used morphological algorithms and filters for discrimination. They

just showed the differentiation between three species-Pilchard Sardine (ma-iwashi), Japanese Horse Mackerel (ma-aji), and Common Mackerel (ma-saba) with an accuracy of 90%, 88% and 90% respectively. Their method is dependent on the size and shape of the morphological filters used.

Junguk Cho et al. [9] used Haar classifiers with the Scythe Butterfly fish as their test species. Their method is heavily dependent on the background environment in the image of the fish and the angle from which the photograph is taken.

Andrew Rova et al. [10] attempt to classify fishes using the method of warping the images prior to classification using SVM on a dataset of 320 images achieving an accuracy of 90%.

Chomptip Pornpanomchai et al. [11] proposed and developed a fish recognition system based on shape and texture. They compared Artificial Neural Networks and Euclidean Distance Method (EDM) working on a test set of 300 images and a training set of 600 images. They achieved an accuracy of 81.67% and 99.00% with EDM and ANN respectively.

Rodrigues et al. [13] recommended an algorithm based on SIFT feature extraction & Principal Component Analysis (PCA) but they worked on a very small dataset of 162 images encompassing 6 different species getting an accuracy of 92%.

S. Sclaroff et al. [14] performed object deformable shape detection and object detection was done through model-based region grouping. Computational complexity is the major drawback of this method.

C. Spampinato et al. [15] worked on 20 underwater videos to detect, track and count fishes with an accuracy of 85%. They performed detection using Moving Average Detection Algorithm and Adaptive Gaussian Mixture Model.

Andres Hernandez-Serna et al. [16] used Artificial Neural Networks for the automatic identification of species. They worked on a dataset of 697 images achieving an accuracy of 91.65%.

Arjun Kumar Joginpelly et al. [17] propose an automatic technique using Gabor filters to extract important features from two species, Epinephelus morio and Ocyurus chrysurus. The proposed algorithm is tested on 200 frames, each containing many fish and non-fish regions. The accuracy is 70.6% for Epinephelus morio and 80.3% for Ocyurus chrysurus.

Deokjin Joo et al. [18] extracted stripe and color patterns of wild cichlids and used Random Forests and SVM for classification achieving an accuracy of 72% on a dataset of 594 wild cichlids. They have a low accuracy and they just target Cichlid fishes.

Yi-Haur Shiau et al. [19] proposed a method of sparse representation-based classification for the recognition and verification of fishes which maximizes the probability of partial rankings thus obtained. They worked on a dataset of 1000 images and achieved the highest accuracy of 81.8% for a particular feature space dimensionality.

S.O. Ogunlana et al. [20] classified using Support Vector Machine technique based on the shape features of the fish. They worked on a training data of 76 fish and testing data of 74 fish achieving an accuracy of 78.59%.

S. Cadieux et al. [21], generates the contours of fish in an unconstrained environment by deploying an infrared silhouette sensor which acquires contours of fish in a constrained flow. When the inputs are noisy, these features give a poor performance. The system has reported classification accuracy around 78%.

D. J. Lee et al. [22], removed edge noise and redundant data points through the development of a shape analysis algorithm. Critical landmark points were located using an algorithm of curvature function analysis. A group of nine fish species was used to test this method. The dataset that was used to perform the experimentation consisted of only 22 images.

M. Nery et al. [23], proposed a methodology based on feature selection. This method develops a feature vector by utilizing a set of descriptors obtained by analysis of the characteristic contribution of an individual descriptor to the overall performance for classification. A classification accuracy of about 85% is reported.

III. PROPOSED METHODOLOGY

Here, we present a methodology for the discrimination of fish species. The dataset used for the concerned work is taken from [24]. The initial step taken by the system aims at removing the noise in the dataset. Application of Image Processing before the training step helps to remove the underwater obstacles, dirt and non-fish bodies in the images. The second step uses Deep Learning approach by implementation of Convolutional Neural Networks(CNN) for the classification of the Fish Species.

In order to get the best results for feature identification and training of the CNN, it is important to provide input image with enhanced features as training sample. The pre-processing consists of the following steps,

```
kernel = np.ones((3,3),np.uint8)
opening = cv2.morphologyEx(thresh,cv2.MORPH_OPEN,kernel, iterations = 2)
sure_bg = cv2.dilate(opening,kernel,iterations=3)
dist_transform = cv2.distanceTransform(opening,0,5)

ret, sure_fg = cv2.threshold(dist_transform,0.7*dist_transform.max(),255,0)
sure_fg = np.uint8(sure_fg)
unknown = cv2.subtract(sure_bg,sure_fg)
```

With the implementation of Otsu's thresholding, a grey level histogram is created from the grayscale image for noise removal, if a pixel of the grayscale is greater than the threshold value, it is considered to be white, else declared as black. The image provides a sure foreground with the fish in focus.

Next step deals with the implementation of Morphological Operations, viz-a-viz, Erosion and Dilation of the binarized image.

In the Erosion of the image, a kernel, precisely a fixed size matrix is convolved over the image. A pixel in the processed image will be taken as 1 only if all the pixels under the kernel are 1, otherwise, the pixel is eroded (0).

Thus, in this step, the thickness of the foreground object(fish) decreases.

The step of Dilation implements the algorithm that a pixel in the processed image is 1 if at least one pixel under the kernel is 1.

Thus, this step is used to join the broken parts of the image from the noise removal step.



Figure 1. Pre-processing with Otsu's binarization, Dilation and Erosion.

The Second step of the procedure is the implementation of a Convolutional Neural Network(CNN) for classification of Fish species.

The input layer of the network takes the 100x100x3 original RGB image stacked with the 100x100x1 image which is the output of the pre-processing stage, thus making the input of 100x100x4, the fully-connected layer where we get the trained output and the intermediate hidden layers. The network has a series of convolutional and pooling layers. Neurons in layer say, 'm' are connected to a subset of neurons from the previous layer of (m-1), where the (m-1) layered neurons have contiguous receptive fields, as shown in Fig(2a).

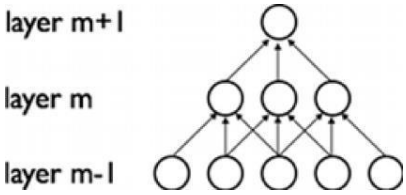


Figure 2a. Graphical flow of layers showing the connection between layers.

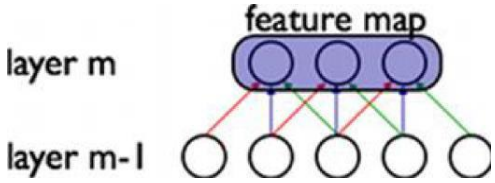


Figure 2b. Graphical flow of layers showing sharing of Weights.

Fig (2b) represents three hidden units. The weights of similar color are shared, thus are inferred to be identical.

The summation of the gradients of the parameters that are being shared results in the gradient of the shared weight. Such similarity thus allows detection of features regardless of their positions in the visual field. In addition to this, weight sharing tends to decrease the number of free learning parameters. Due to this control, CNN tends to achieve better generalization on vision problems.

The Max-pooling layers act as non-linear down sampling, in which the input image is partitioned into non-overlapping rectangles. The output of each sub-region is the maximum value.

The Convolution Layer is the first layer of the CNN network. The structure of this layer is shown in fig (3). It consists of a convolution mask, bias terms and a function expression. Together, these generate the output of the layer. The figure below shows a 5x5x4 mask that performs convolution over a 100x100x4 input feature map. Upon application of 32 such 5x5x4 filters, the resultant output is a 96x96x32 matrix.

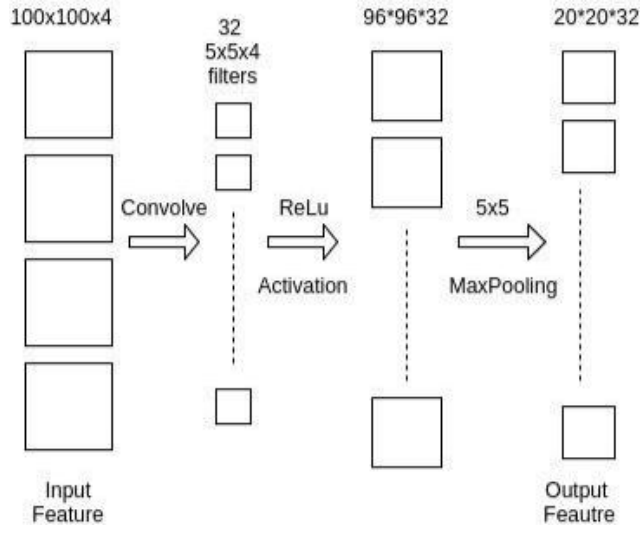


Figure 3. Processing the input feature with 32 filters and max-pooling.

The next layer in the network is a subsampling layer. The Subsampling layer is designed to have the same number of planes as the convolution layer. The purpose of this layer is to reduce the size of the feature map. It divides the image into

blocks of 5x5 and performs max-pooling. Sub-sampling layer preserves the relative information between features and not the exact relation.

Figure 3 shows how the input features are processed with 32 filters and max-pooling. The above process is repeated two times again once with 64 filters and then with 32 filters.
The final output is connected to a fully connected layer which is further connected to an 80% dropout layer and lastly another fully connected layer which classifies the images into appropriate categories.

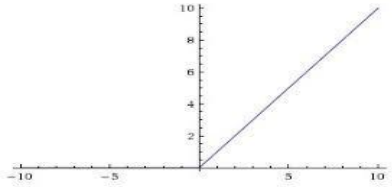
The aim of the training algorithm is to train a network such that the error is minimized between the network output and the desired output.
In the proposed method, we provide the comparison of different Activation Functions that will be applied to the different Layers in the CNN. The following are the three Activation functions, namely,

a)ReLU b) Sigmoid c) tanh

Mathematical definition of the above functions explained below:

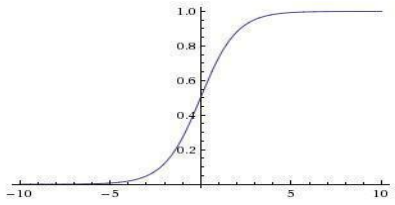
a) **ReLU** : Rectified Linear Unit is :

ReLU: h = max (0, a), where a = W*x+b



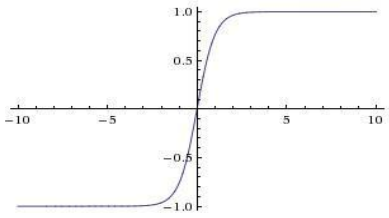
b) **Softmax**:

$\sigma(z) = (e^{z} / \sum^{k}_{k=1} (e^{z}_{k}))$



c) **tanh**: Mathematical definition states,

$\tanh = (e^{+x} + e^{-x}) / (e^{-x} - e^{+x})$



The loss function used after the fully-connected layer is Cross-entropy, Mathematically,

$$Hy'(y) := -\sum i(y'i \log(yi) + (1-y'i)\log(1-yi))$$

which can be explained as a (minus) log-likelihood for the data y'_i, under a model y_i.

TABLE 1. ALGORITHM FOR PRE-PROCESSING AND CONVOLUTIONAL NETWORK:

Algorithm 1. Fish Classification Algorithm

Input:

- Training set $X = \{(X_i, c_i)\}^{N}_{i=1}$ from the given dataset undergoes Otsu's binarization and thresholding to create a grey-level histogram, further implementing Dilation, Erosion and Subtraction to remove the background noise and other water objects.
- Pyramid Mean Shifting implemented to highlight the outlines of the fish object in the image with the kernel function $k(x_i - x)$.

Processing:

- Architect the model for efficient mapping of features and set appropriate hyper-parameters with the input dimensions as that of X.
- Compute Feature matrices of each layer which gets more detailed with further layers and feed it as input to next layer.
- Feed the Fully Connected Layer with the final feature matrix. Repeat the previous step with the three Activation functions discussed, ReLU, tanh, Sigmoid.

Output:

- Using *Adam optimizer* for the Fully Connected Layer outputs

IV. EXPERIMENTAL RESULTS

The proposed method was tested in Python on the dataset Fish4Knowledge [24] of 27,142 images shown in Figure 3 below.





Figure 2. Dataset used

Table 2 provides insights into the number of samples used for each different species covered. Table 3 shows the results as the accuracy of the correctly predicted test images in the sample as given by Eqn. (5) shown below:

$$Accuracy = \frac{TP}{TP + FP} \times 100\% \quad \text{- (5)}$$

Species	Number of Images
Plectroglyphidodon dickii	11312
Chromis chrysura	2683
Amphirion clarkii	3593
Chaetodon lunulatus	4049
Chaetodon trifascialls	2534
Myripristis kuntee	190
Acanthurus nigrofuscus	450
Hemigymnus fasciatus	218
Chaetodon trifascialis	242
Neoniphon sammara	298
Abudefduf vaigiensis	198
Canthigster valentini	148
Pomocentrus molucensis	180
Hemigymnus fasciatus	190
Scolopsis billineate	142
Neoniphon sammara	206
Scaridae	149
Pemphereis vanicolensis	156
Zanclus cornutus	129
Balistapus undulatus	221
Zebrasoma scopas	116
Total	**27,142**

Table 2. The Species used in the Dataset

Activation Function Used	Overall Accuracy
ReLU	**96.29%**
tanh	**72.62%**
Softmax	**61.91%**

TABLE 3. RESULTS

V. CONCLUSION AND FUTURE WORK

The proposed method of the classification of fish species gives an accuracy of 96.29% which is very high compared with the other current implemented methods used for this application. Hence the proposed approach can certainly be used for real time applications as the computation time is 0.00183 seconds per frame. The method couldn't achieve 100% accuracy as some images couldn't be classified accurately due to the effect of background noise and other water bodies. We plan to improvise our algorithm further by implementing Image Enhancement techniques to counter for the lost features in the images.

VI. REFERENCES

[1] White, D. J., Svellingen, C., Strachan, N. J. C., Automated measurement of species and length of fish by computer vision, Elsevier, 2006.
[2] F. Storbeck and B. Daan, Fish species recognition using computer vision and a neural network, Elsevier, 2000.
[3] Nery, M. S., Machado, A. M., Campos, M. F. M., Padua, F. L. C., Carceroni, R., Queiroz-Neto, J. P, Determining the appropriate feature set for fish classification tasks, Proceeding of the XVIII Brazilian Symposium on Computer Graphics and Image Processing, 2005.
[4] M. C.Chuang, J. N.Hwang, C. Rose, Aggregated Segmentation of Fish from Conveyor Belt Videos, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1807-1811, 2013.
[5] N. Castignolles, M. Catteon, M. Larinier, Identification and counting of live fish by image analysis, SPIE. Vol 2182, Image and Video Processing II, 1994.
[6] K.A. Mutasem, B.O. Khairuddin, N. Shahrulazman and A. Ibrahim (2010). "Fish Recognition Based on Robust Features Extraction from Size and Shape Measurements Using Neural Network". Journal Computer Science, Volume 6, Issue 10
[7] C. Spampinato, D. Giordano, R. D. Salvo, Y-H C-Burger, R.B. Fisher, G. Nadarajan, Automatic Fish Classification for Underwater Species Behaviour Understanding
[8] Y. Nagashima, Takakazu Ishimatsu, A Morphological Approach to Fish Discrimination, IAPR Workshop on Machine Vision APPLICATIONS, Nov. 17-19,1998
[9] B. Benson, J. Cho, D. Goshorn, R. Kastner, Field Programmable Gate Array (FPGA) Based Fish Detection Using Haar Classifiers, American Academy of Underwater Sciences, March 1, 2009.
[10] A. Rova, G. Mori, L. M. Dill, One fish, two fish, butterfish, trumpeter: Recognising fish in underwater videos, IAPR Conference on Machine Vision Applications

[11] C. Pornpanomchai, B. Lurstwut, P. Leerasakultham, W. Kitiyanan, Shape and Texture based fish image recognition system, Kasetsart J. (Nat. Sci.) 47 : 624 - 634 (2013)
[12] Mehrtash T.Harandi, Conrad Sanderson, Sareh Shirazi, etc. Graph Embedding Discriminant Analysis on Grassmannian Manifolds for Improved Image Set Matching. CVPR2011.2705-2712,2011
[13] Rodrigues, M.T.A., Freitas, M.H.G., Pádua, F.L.C. et al. Pattern Anal Applic (2015) 18: 783. doi:10.1007/s10044-013-0362-6
[14] S. Sclaroff, L. Liu, Deformable Shape Detection and Description via Model-Based Region Grouping, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 23, NO. 5, MAY 2001
[15] Spampinato, C, Chen-Burger, Y-H, Nadarajan, G & Fisher, R 2008, 'Detecting, Tracking and Counting Fish in Low Quality Unconstrained Underwater Videos'. inProc. 3rd Int. Conf. on Computer Vision Theory and Applications (VISAPP). vol. 2, pp. 514-519.
[16] Hernández-Serna, Andrés, and Luz Fernanda Jiménez-Segura. "Automatic Identification of Species with Neural Networks." Ed. Mark Costello. PeerJ 2 (2014): e563. PMC. Web. 30 Sept. 2016.
[17] AK Joginipelly, D Charalampidis, G Ioup, J Ioup, Ch Thompson, Species-Specific Fish Feature Extraction Using Gabor Filters
[18] Joo D1, Kwan YS, Song J, Pinho C, Hey J, Won YJ, Identification of Cichlid Fishes from Lake Malawi using Computer Vision, 2013 Oct 25;8(10):e77686. doi: 10.1371/journal.pone.0077686
[19] Y. H Shiau, F-P. Lin, C-C Chen, Using Sparse Representation for Fish Recognition and Verification in Real World Observation, ICONIP;12 Proceedings of the 19th International Conference on Neural Information Processing-Volume Part IV, pages 75-82
[20] S.O. gunlana, O. Olabode, S.A.A. Oluwadare, G.B. Iwasokun, Fish Classification Using Support Vector Machine, African Journal of Computing & ICT, Vol 8. No. 2 June, 2015
[21] S. Cadieux, F. Lalonde, and F. Michaud, "Intelligent System for Automated Fish Sorting and Counting," IEEE IROS, pp. 1279–1284, 2000.
[22] D. J. Lee, S. Redd, R. Schoenberger, X. Xiaoqian, and Z. Pengcheng, "An Automated Fish Species Classification and Migration Monitoring System," in Conf. of the IEEE Industrial Electronics Society, 2003, pp. 1080–1085.
[23] M. Nery, A. Machado, M. Campos, F. Padua, R. Carceroni, and J. Queiroz-Neto, "Determining the Appropriate Feature Set for Effective Fish Classification Tasks," in SIBGRAPI, 2005, pp. 173–180.
[24] B. J. Boom, P. X. Huang, J. He, R. B. Fisher, "Supporting Ground-Truth annotation of image datasets using clustering", 21st Int. Conf. on Pattern Recognition (ICPR), 2012.
[25] T. Wang and P. Shi. Kernel grassmannian distances and discriminant analysis for face recognition from image sets. Pattern Recognition Letters, 30(13):1161–1165, 2009