SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(29th Edition)

MAY 2015

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 **Twenty Ninth** 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 2015**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A Complete Method Configuration Process for Configuring Project-Specific Methods, ***3.Rinky Dwivedi*** and ***3.Daya Gupta***, DTU

2. A Novel Wavelet Modulation Scheme for Single Phase Inverter, Ankit Yadav and ***3.Dr. M. M. Tripathi***, Department of Electrical Engineering, DTU

3. Adiabatic Vedic Multipler Design Using Chinese Abacus Approach, ***6.Diksha Ruhela*** and ***3.Dr. Malti Bansal***, Electronics & Communication Engineering, DTU

4. Affinity of Smectite and Divalent Metal Ions (Mg2+, Ca2+, Cu2+) with L-leucine: An Experimental and Theoretical Approach Relevant to Astrobiology, Pramod Pandey1 & Chandra Kala Pant, Kavita Gururani, Priyanka Arora, Neetu Pandey, Preeti Bhatt, Yogesh Sharma, Jagmohan Singh Negi and ***3.Mohan Singh Mehata***, Department of Applied Physics, DTU

5. Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behavior, ***3.Ajay Kumar1, Neha Gupta*** and ***3.Rishu Chaujar***, Microelectronics Research Lab, Department of Engineering Physics, DTU

6. Analytical Level Crossing Rate and Average Outage Duration for MRC Combiner in Log Normal Shadowed Channels, ***3.Rupender Singh*** and ***3.Sanjay Kumar Soni***, Department of Electronics & Communication, DTU

7. Analytical study: Workplace Environment in IT Companies and Ranking of Companies using AHP –Topsis Method, Bhavna Saighal, M. K. Bhat and ***6.Anjana Gupta***, Department of Mathematics, DTU

8. Ant Colony Optimization Based Orthogonal Directional Proactive–Reactive Routing Protocol for Wireless Sensor Networks, ***7.1Aarti Jain*** and B. V. Ramana Reddy, Electronics and Communication, DTU

9. Assessment of Critical Enablers for Flexible Supply Chain Performance Measurement System Using Fuzzy DEMATEL Approach, Mohit Tyagi, ***1.Pradeep Kumar*** and Dinesh Kumar, DTU

10. Characterisation Study and Correlation Analysis For Water Quality of Dal Lake, India, J.N.Sharma, ***3.Raj Shailesh Kanakiya*** and ***3.S.K.Singh***, Department of Environmental Engineering, DTU

11. Consequence of pH variation on the dielectric properties of Cr-doped lithium ferrite nanoparticles synthesized by the sol–gel method, Manish Srivastava, R.K. Mishra, ***3.Jay Singh***, Neha Srivastava, Nam Hoon Kim and Joong Hee Lee, Department of Applied Chemistry & Polymer Technology, DTU

12. Adiabatic Vedic Multipler Design Using Chinese Abacus Approach, ***7.Diksha Ruhela*** and ***3.Dr. Malti Bansal***, Electronics & Communication Engineering, DTU

13. Credit Card Fraud Detection Using Advanced Combination Heuristic and Bayes' Theorem, ***8.Sahil Hak, 8.Suraj Singh*** and ***8.Varun Purohit***, Computer Engineering, DTU

14. Effect of Single and Multiple-Pass Friction Stir Processing on Microstructure, Hardness and Tensile Properties of a 99.99% Cu with Carbon Nano Tubes, ***3.V. Jeganathan Arulmoni, 3.Ranganath M. S.*** and ***3.R. S. Mishra,*** Mechanical Engineering, DTU

15. Experimental investigation into tool aging effect in ball end magnetorheological finishing, ***3.Mahendra Singh Niranjan*** and Sunil Jha, DTU

16. Fingerprint denoising using ridge orientation based clustered dictionaries, Kuldeep Singha, ***3.Rajiv Kapoor*** and ***3.Raunaq Nayar***, Electronics & Communication, DTU

17. Fuzzy Modelling For Selection Of Most Economical Green Building Out Of N-Alternatives, Sunita Bansal, Prof. (Dr.) Srijit Biswas, 3.Prof. (Dr.) S. K. Singh and Sunita Bansal, Department of Environmental Engineering, DTU

18. GIS Application in Urban Traffic Air Pollution Exposure Study: A Research Review, ***3.Amrit Kumar, 3.Rajeev Kumar Mishra*** and ***3.S. K. Singh***, Environmental Engineering, DTU

19. Hybrid Classifier based Human Activity Recognition using the silhouette and cells, ***3.D. K. Vishwakarma*** and ***3.Rajiv Kapoor***, Electronics & Communication Engineering, DTU

20. Modified Hysteresis Switching Scheme for 3-Phase Voltage Source Inverter, Ankit Yadav and ***3. Dr. M. M. Tripathi***, Department of Electrical Engineering, DTU

21. Nanomechanical Properties of Cement Paste Containing Silica Fume, Salim Barbhuiya, PengLoy Chow1 and ***3.Arnesh Das***, DTU

22. Use of ecofriendly in Low Temperature Circuit in Terms of First Law and Second Law Efficiency of Three Stages Cascade Refrigeration of Biomedical Applications, ***3.R.S. Mishra,*** Mechanical Engineering Department, DTU

23. Sustainability analysis of eight ecofriendly refrigerants two-stage vapour compression refrigeration system-using energy-exergy performance, Kapil Chopra, V. Sahni and ***3.R.S. Mishra***, Mechanical Engineering, DTU

24. Novel back‐contact back‐junction SiGe (BC‐BJ SiGe) solar cell for improved power conversion efficiency, ***Rahul Pandey*** and ***Rishu Chaujar***, Microelectronics Research Lab, Department of Engineering Physics, DTU

25. On the Assessment of Segmentation Methods for Images of Mosaics, Gianfranco Fenu, ***3.Nikita Jain*** and Eric Medvet1, Software Engineering, DTU

26. Organic–Inorganic Hybrid Nanocomposite-Based Gas Sensors for Environmental Monitoring, Ajeet Kaushik, Rajesh Kumar, Sunil K. Arya, Madhavan Nair and ***3.B. D. Malhotra*** and Shekhar Bhansali, Biotechnology, DTU

27. Oxide bound impact on hot-carrier degradation for gate electrode workfunction engineered (GEWE) silicon nanowire MOSFET, ***Neha Gupta, Ajay Kumar*** and ***Rishu Chaujar***, Microelectronics Research Lab, Department of Engineering Physics, DTU

28. PHOSPHATE SOLUBILIZING PLANT GROWTH PROMOTING MICROBES, Faria Fatima, Neelam Pathak, Preeti Bajpai and ***3.Smita Rastogi Verma***, Department of Biotechnology, DTU

29. Physiological responses of Catla catla larvae fed with Achyranthes aspera seed enriched diet and exposed to UV-B radiation, ***3.Jai Gopal Sharma***, Moirangthem Kameshwor singh and Rina Chakrabarti, Biotechnology, DTU

30. Promoting Human Resources Skill Developments and Impowering India with Solar Power Technologies, Education, Training and Innovation, Dr. L P Singh, ***3.Dr. J P Kesari***, Dr. A K Ojha and Mr. Gaurav Kumar Gaur, DTU

31. Shear Behaviour of Rock Joints Under CNL and CNS Boundary Conditions, ***3.Amit Kumar Shrivastava,*** K. Seshagiri Rao, Civil Engineering, DTU

32. Software Reliability Modeling using Soft Computing Techniques: Critical Review, Kaswan KS, Choudhary S and ***3.Sharma K***, Computer Science and Engineering, DTU

33. SUSTAINABLE HANDLING OF CONSTRUCTION AND DEMOLITION (C & D) WASTE, 6.Shishir Bansal and 3.S K Singh, Environmental Engineering Department, DTU

34. TENSILE BEHAVIOUR OF 3-PLY LAMINATE COMPOSITE OF SHEET METALS, ***3.Vijay Gautam*** and ***6.Bijender Prasad***, Mechanical Engineering, DTU

35. Text Watermarking using Sinusoidal Greyscale Variations of Font based on Alphabet Count, ***3.Jeebananda Panda, 8.Nishant Gupta, 8.Parag Saxena, 8.Shubham Agrawal, 8.Surabhi Jain*** and ***3.Asok Bhattacharyya***, ECE, DTU

36. Theoretical Investigation of Solar Energy Driven Combined Power and Refrigeration Cycle, ***3.Rahul Singh, 3.Rajesh Kumar, 3.Supriya Vats*** and ***3.Pavan Kumar Yadav***, Department of Mechanical Engineering, DTU

37. Thermal Evaporation- Modeling and Microstructure Studies of Indium and Tin Deposition, Shailendra Kumar Gaur, R. S. Mishra, Department of Mechanical, Production & Industrial and Automobiles Engineering, DTU

38. Using Application of Dezert- Smarandache Theory on a new framework for protecting MANET, Ekata Gupta ***3.Dr.S.K.Saxena***, CSE Department, DTU

39. Voltage differencing transconductance amplifier based resistorless and electronically tunable wave active filter, ***3.Neeta Pandey***, Praveen Kumar and Sajal K. Paul, Electronics and Communication, DTU

1. Vice Chancellor	***1.1. Ex Vice Chancellor***
2. Pro Vice Chancellor	***2.1. Ex Pro Vice Chancellor***
3. Faculty	***3.1. Ex Faculty***
4. Teaching-cum-Research Fellow	
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***

A Novel Wavelet Modulation Scheme for Single Phase Inverter

Ankit Yadav
Department of Electrical Engineering
Institute of Engineering &Technology (JKLU)
Jaipur, India
ankit.jnvf@gmail.com

Dr. M. M. Tripathi
Department of Electrical Engineering
Delhi Technological University
New Delhi, India
mmtripathi@dce.ac.in

***Abstract*-- DC-AC Inverters are key components in various industrial applications that include power supplies, motor drive systems and power systems. Now-a-days dc-ac inverters (both 1-phase & 3-phase) are finding more attention in renewable energy systems also. In this paper a new kind of switching scheme is discussed where gate pulses for switching of Single Phase 4 pulse voltage source (VS) DC-AC inverters are generated using wavelet modulation. Simulation for different loading conditions is performed and results are compared with results of PWM and Square pulse switching under same loading conditions. The wavelet modulation technique has shown tremendous advantages over PWM technique.**

***Index Terms*--** PWM, Wavelet, Square Pulse, Scaling, 1-ph 4-pulse H-bridge voltage source DC-AC inverter.

I. INTRODUCTION

DC-AC inverters have become the key component of the power system network in last few years. With the increase in number of DC-AC inverters in power system, better operation of inverters has been emerged as priority for power system engineers. The operation of these inverters are controlled by gate signals and thus the output, harmonic, efficiency and other factors can be controlled by controlling gate signal [1], [2].

A proper switching strategy can give an efficient and better control over output of DC-AC inverters. Several factors of AC-DC inverter operation that must be taken care while developing the switching strategy are as follows [1]-[3] .

1) Output RMS voltage and current.
2) Output peak voltage and current
3) Total harmonic distortion (THD).
4) Switching losses.
5) Practical implementation.

In Past decade several switching strategies have been developed to tackle the issue. Square pulse switching, pulse width modulation (PWM), random PWM (RPWM), selected harmonic elimination (SHE), delta modulation, hysteresis & variable band, optimal PWM etc. are the few well known switching techniques.

In this paper a new switching technique called as wavelet modulation has been discussed. A wavelet is a wave–like oscillation with amplitude that begins at zero, increases, and then decreases back to zero. It can typically be visualized as a brief oscillation like one recorded by a seismograph or heart monitor [4]. In this paper the gate pulse for 1-ph 4-pulse H-bridge voltage source DC-AC inverter has been generated using the basic idea of wavelet modulation [3], [5]-[7]. The simulations have been performed on various loads and result of wavelet modulation is compared with other switching techniques. The performance of wavelet modulation in DC-AC inverter switching is very promising in terms of low THD and high output compared to other conventional techniques.

The paper has been organized in six sections. Section II presents the overview of modulation techniques. Section III discusses the wavelet modulation technique. Section IV shows the simulations model used. Results of simulation are presented in Section V. Conclusion and future work is mentioned in section VI.

II. MODULATION OVERVIEW

The wavelet modulation for DC-AC 1-phase inverter is a process of sampling and reconstruction of signal. The various elements of modulation are discussed below.

A. Reference Signal

The output of DC-AC 1-phase inverter is sinusoidal wave. In wavelet modulation the main focus is on generating samples of a sinusoidal reference signal and reconstructing the sinusoidal output from these samples. The reference signal (S_M) of frequency f_M is as in [5]:

$$S_M = Sin(2\pi f_M t) \quad (1)$$

$$f_M = \frac{1}{T_M} \quad (2)$$

Samples are generated from this reference signal over a time period of T_M. Each two samples corresponds to one group of samples (i.e. each group of sample of have two samples) denoted by s. Thus s is the total number of group of samples over time period T.

B. Output Waveform

The output current waveform of DC-AC 1-phase voltage source inverter is desired to be sinusoidal wave. It has been

observed in case of PWM, as shown in fig. 1, that the rectangular output voltage pulses of inverter gives the sinusoidal current waveform and the quality of the output current waveform depends upon the width of rectangular output voltage pulses as well as their number.

These output pulses are function of gate pulses resulting from the switching logic which could be PWM, RPWM, SHE, Square pulse or the one used in this paper i. e. wavelet modulation.

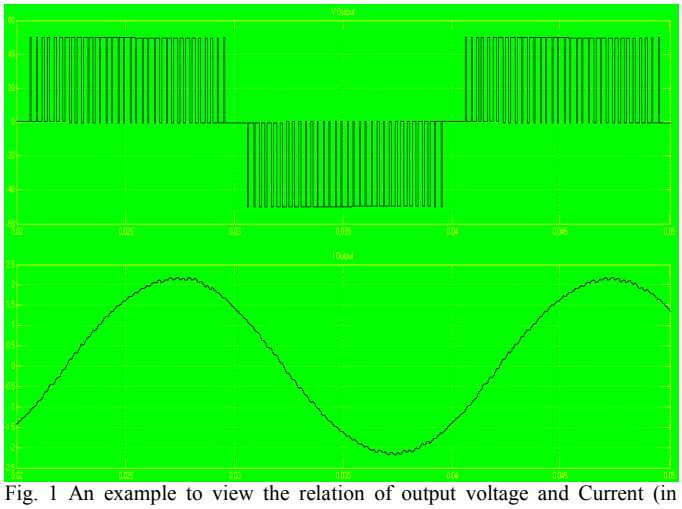


Fig. 1 An example to view the relation of output voltage and Current (in PWM method)

C. *Sampling and Reconstruction*

Sampling and Reconstruction is the link between the reference signal and the output waveform. As the nature of output waveform can be controlled by varying the width of rectangular pulses and so the sampling and reconstruction also needed to be non-uniform in nature. Thus only a non-uniform sampling can lead to sinusoidal output. The gate pulses of inverter are basically ON-OFF signal and the ON signal can be modeled from Haar scaling function [6]. Duration and location of the pulses can be controlled by suitable scaling function.

III. WAVELET MODULATION

A wavelet modulation (wavelet transformation) starts with defining wavelet function □ (mother wavelet) and scaling function □ (father wavelet) and using these functions a new scaling function can be derived as per requirement. The inverter gate pulses are basically ON-OFF pulses, so Haar's wavelet and scaling function can be used for deriving a scaling function which can be used further for switching of inverter.

The Haar wavelet function is defined as in (3) [8], [9].

$$(t) = \begin{cases} 1, 0 \le t \le \frac{1}{2} \\ -1, \frac{1}{2} \le t \le 1 \\ 0, \ otherwise \end{cases} \qquad (3)$$

The Haar wavelet function can be represent as in fig. 2 Haar Scaling function is defined as in (4).

$$\phi(t) = \begin{cases} 1, 0 \le t \le 1 \\ 0, \ \text{otherwise} \end{cases} \qquad (4)$$

Haar Scaling function can be represent as in fig. 3

Further the wavelet modulation involves multi-resolution analysis (MRA) which is analysis of the space $L^2(R)$ consist of a sequence of nested subspaces as in (5) [10].

$$V_0 \subset V_1 \subset V_2 \subset V_3 \subset V_4 \ldots\ldots\ldots\ldots \subset L^2(R) \qquad (5)$$

As discussed above using the Haar scaling function a new scaling function has been derived as in (6)

$$_k(t) = (a^{k+1}t) + \left(a^{k+1}(t - 1 + a^{-(k+1)})\right) \qquad (6)$$

Where k = 1, 2, 3…. and the value of 'a' should be chosen according to the frequency of output signal.

The dual scaling function can be obtained by subtracting it from Haar scaling function. The other factors are t_{s1} and t_{s2} which can be defined as in (7) and (8).

$$T_{s1} = d * T_p + 2^{-j-1} \qquad (7)$$

$$T_{s1} = d * T_p + T_p + 2^{-j-1} \qquad (8)$$

Where s = 1, 2, 3, 4 ….. S

At time t_{s1}, ON pulse start and at t_{s2}, it stops.

S is the number of samples in one cycle (i.e. time period Tm).

Tp is the time of one sample and mathematically can be define as Tp = 1/S;

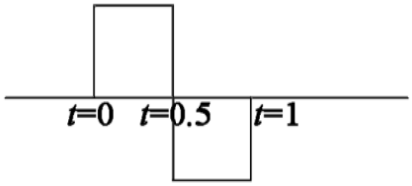


Fig. 2 Haar Wavelet function

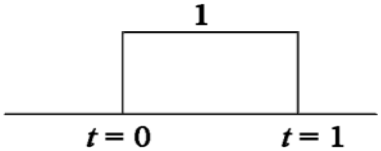


Fig.3 Haar scaling function

IV. MODELLING AND SIMULATION

Simulation has been performed using SIMULINK facility of MATLAB R2011b, on the 1-ph 4-pulse H-bridge voltage source DC-AC inverter. The SIMULINK model of the circuit is shown in fig.4. The wavelet modulation has been implemented using S-function block of SIMULINK [11]. Programming of S-function is done using C language. The inverter is based on 4 switch topology using the IGBT as the switch and controlled by gate signals.

The all four IGBT switches operate according to the gate pulses provided to them. When IGBT.1 and IGBT.2 both are conducting then they generates positive cycle across the load and similarly negative cycle gets generated by operating IGBT.3 and IGBT.4.

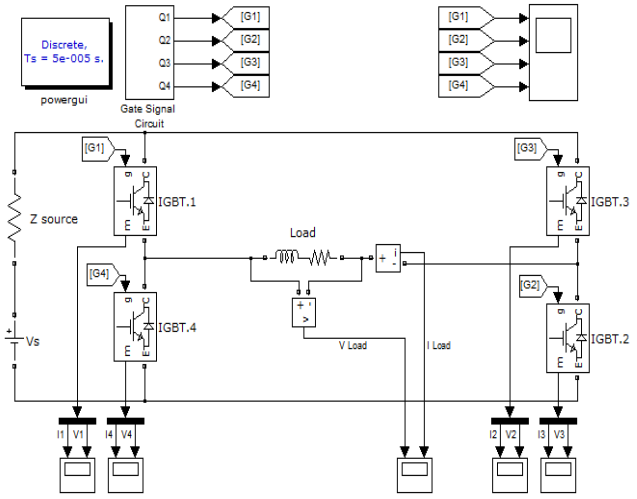


Fig. 4 MATLAB model of 1-ph 4-pulse H-bridge voltage source DC – AC inverter using IGBTs.

V. RESULTS AND DISCUSSION

To analyze the wavelet modulation techniques developed above, simulation has been performed for various kinds of linear and non–linear loads. Resistive load (R load) and resistive–inductive load (R–L load) are used for simulation and testing of new wavelet modulation technique. The simulation has also been performed for PWM and Square pulse techniques under same loading condition. Results of all the three techniques have been used for comparative analysis of performance.

The simulation tests are performed for current output of 50Hz. The loading conditions and other parameter used for simulation are as follows-

1) Source voltage (Vs) = 50V
2) Source impedance (Zs) = 0.1 Ω
3) R Load = 25 Ω
4) R – L load = 15+ j12.566 Ω
5) Modulating index for PWM = 0.85
6) Modulation frequency for PWM = 2KHz

A. R – L Load

Simulation test is performed using resistive – inductive load of given value on the SIMULINK model of 1-ph 4-pulse H-bridge voltage source DC-AC inverter using wavelet modulation by taking 12 groups of samples in each cycle. Similarly simulation has been performed for PWM and square pulse methods. The plots of the results obtained from Simulation are in fig.6 and fig.7 and 8 shows the THD in voltage and current waveforms.

The results obtained from these simulations show that THD of current and voltage in wavelet modulation is better than Square pulse (1 Pulse) switching.

In the case of PWM, THD in load current is less than the same in the wavelet modulation in present case. However it is significant point to observe that the voltage THD and output voltage for wavelet modulation is better than PWM.

A comparative result of performances of all the three methods is also presented in table 1.

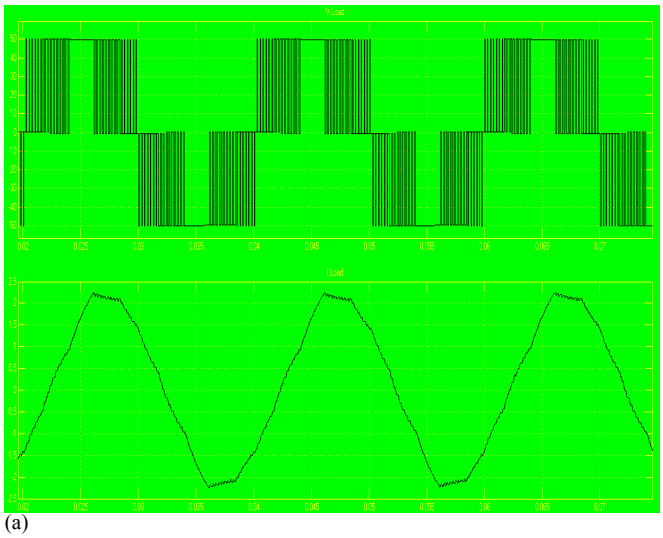
B. R Load

Similar to R–L load, Simulation is also performed with pure resistive load using the SIMULINK model of inverter shown in fig. 4. In this case also simulations have been performed for all the three switching techniques, wavelet modulation, PWM and square pulse. Performance results of these simulation tests are shown in table 1. The waveforms of voltage and current is presented in fig.9 and THD of voltage and current waveform are presented in fig.10 and 11 respectively.

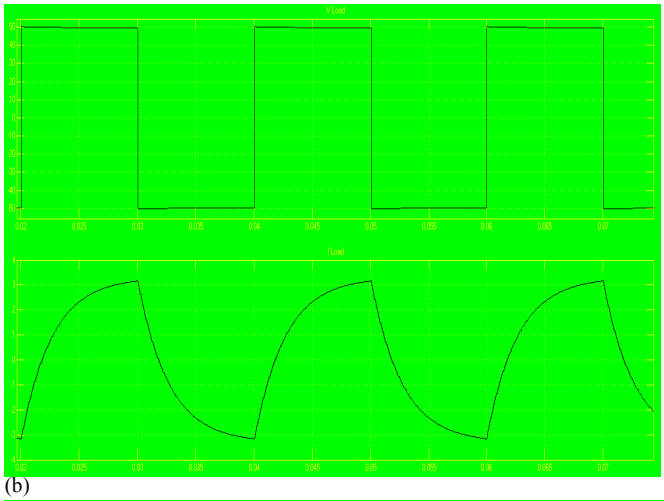
In comparison to PWM and Square wave, wavelet modulation has given much better result. The THD of current as well as of voltage are much lesser in case of wavelet modulation than that in PWM. The load current and voltage have higher value in case of wavelet modulation than that in case of PWM. Wavelet modulation gives improved results in terms of THD than both the PWM and Square pulse switching scheme.

TABLE 1
SIMULATION RESULT FOR VARIOUS LOADS

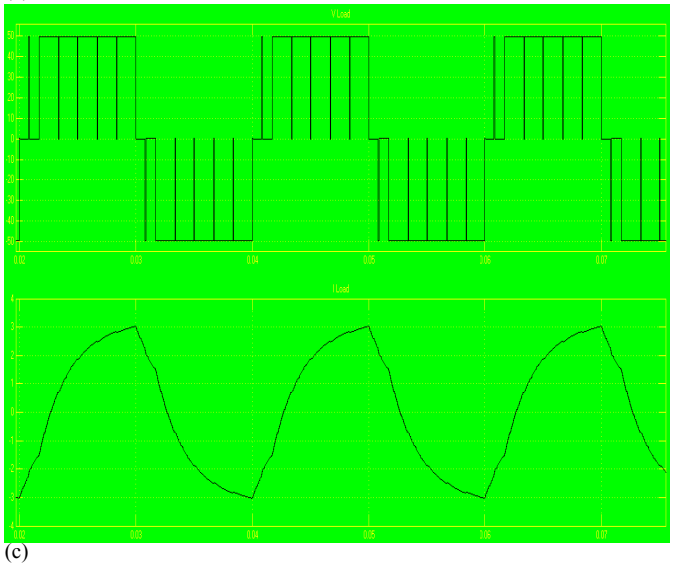
Load Types	Parameters	PWM	WM	1 Pulse
R – L load	Vrms	30.11	41.95	44.83
	Irms	1.539	2.144	2.291
	Vpeak	42.58	59.32	63.41
	Ipeak	2.176	3.032	3.24
	V THD	70.47%	38.63%	48.50%
	I THD	4.76%	12.38%	17.73%
R load	Vrms	30.15	42.06	44.83
	Irms	1.206	1.683	1.793
	Vpeak	42.62	59.49	63.4
	Ipeak	1.705	2.38	2.563
	V THD	69.85%	38.56%	48.34%
	I THD	69.85%	38.56%	48.34%



(a)

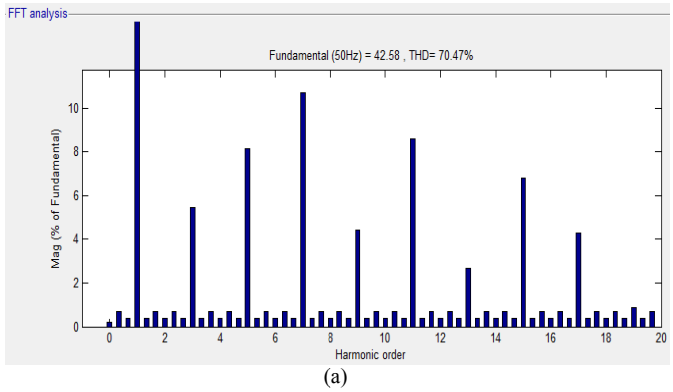


(b)

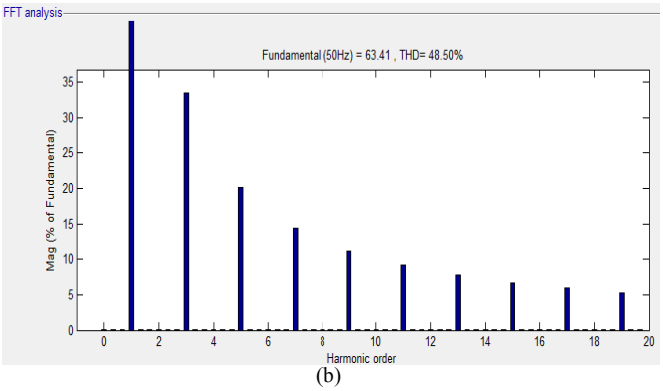


(c)

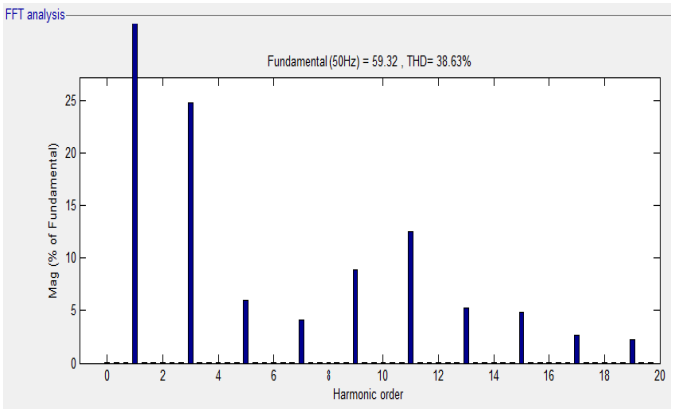
Fig. 6 Voltage and Current output of simulation results for (a) PWM switching, (b) wavelet modulation switching, (c) Square Pulse Switching.



(a)

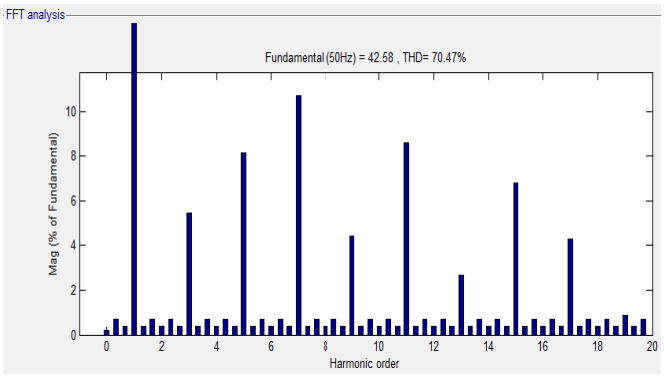


(b)

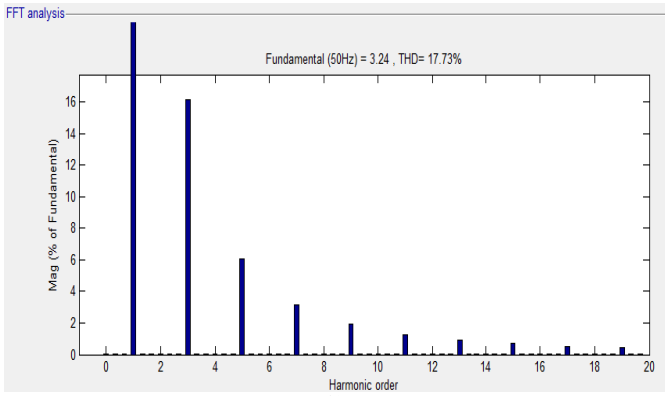


(c)

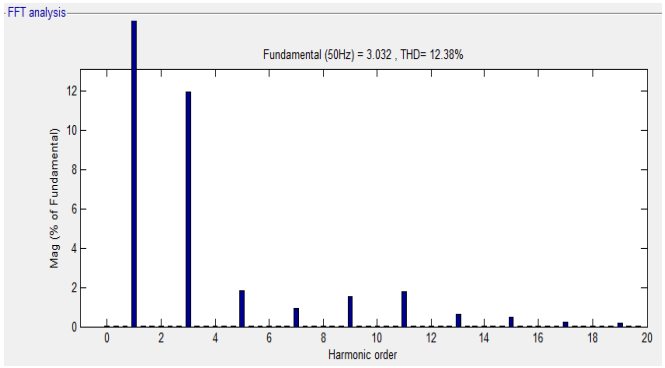
Fig. 7 THD in load voltage waveform of single phase H-bridge VS DC–AC with R-L load (a) PWM (b) square pulse (c) wavelet switching scheme.



(a)

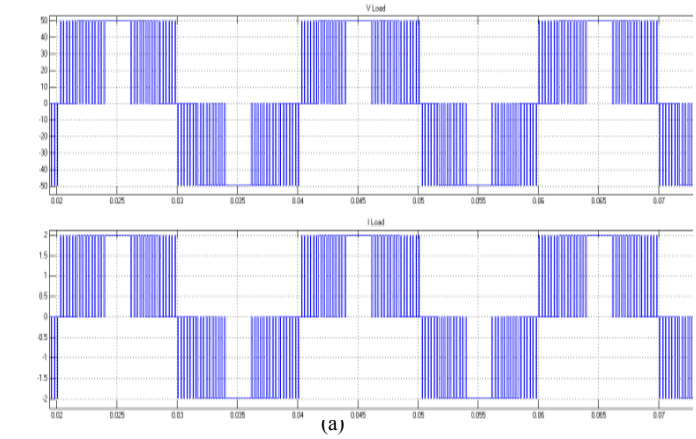


(b)

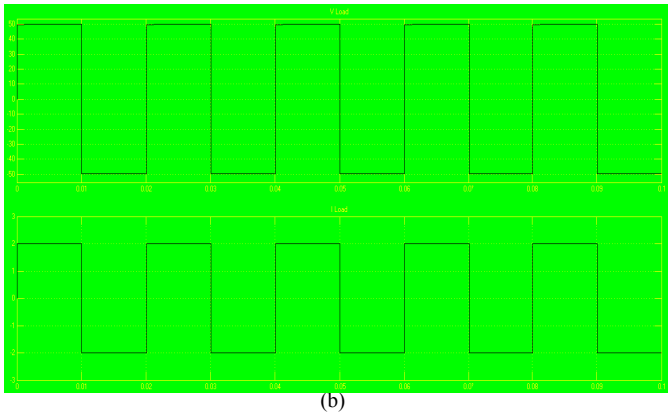


(c)

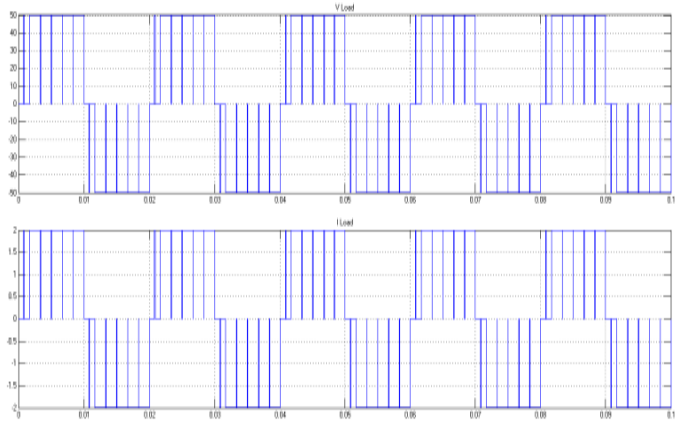
Fig.8 THD in load current waveform of single phase H-bridge VS DC–AC with R-L load (a) PWM (b) square pulse (c) wavelet switching scheme.



(a)

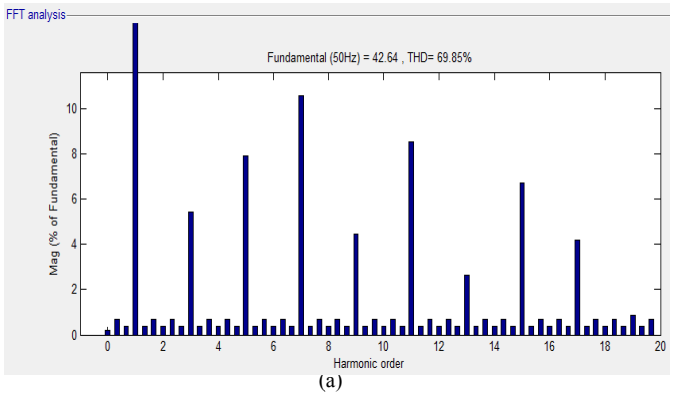


(b)

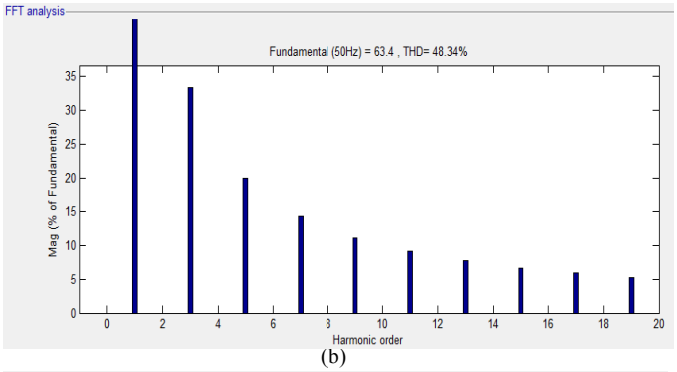


(c)

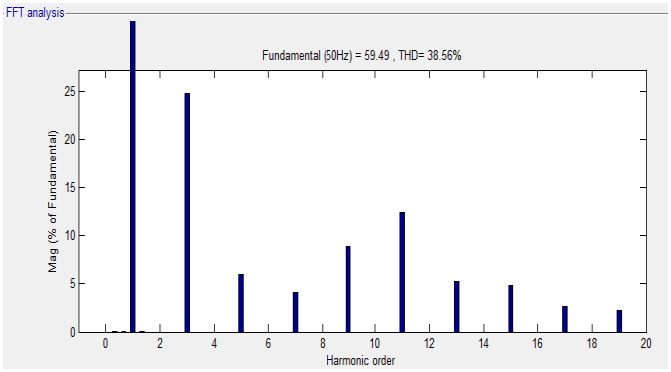
Fig. 9 Voltage and Current output of simulation results for (a) PWM switching, (b) wavelet modulation switching, (c) Square Pulse Switching.



(a)

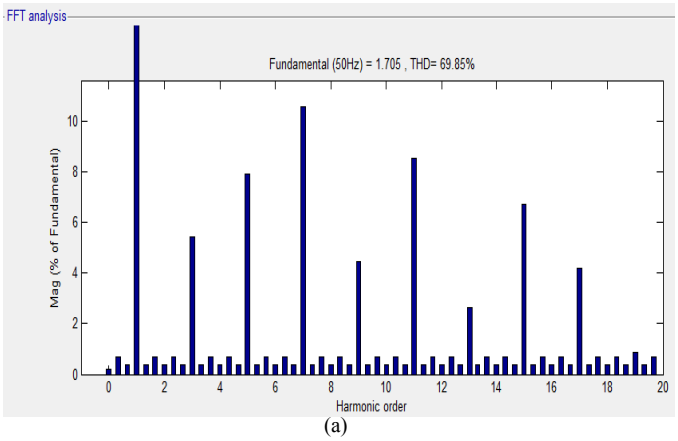


(b)

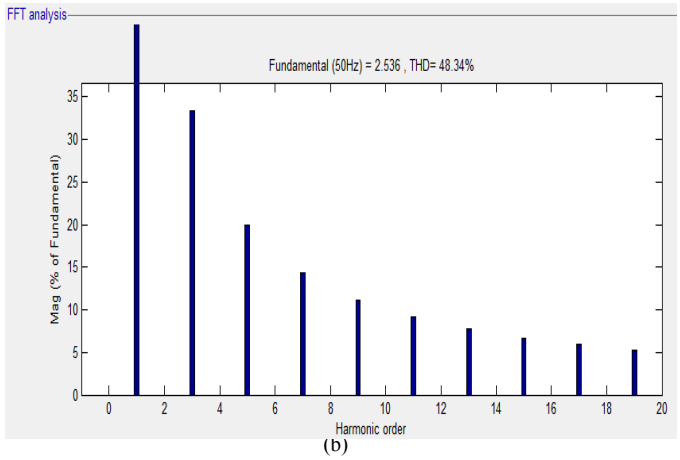


(c)

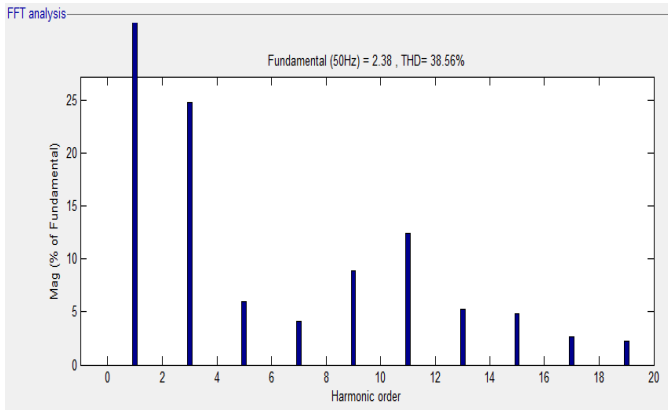
Fig. 10 THD in load voltage waveform of single phase H-bridge VS DC–AC with R-L load (a) PWM (b) square pulse (c) wavelet switching scheme.



(a)



(b)



(c)

Fig. 11 THD in load current waveform of single phase H-bridge VS DC–AC with R-L load (a) PWM (b) square pulse (c) wavelet switching scheme.

VI. Conclusion

. Wavelet modulation technique has given improved results over square pulse switching techniques and it has performed better in terms of Voltage harmonics and output but the wavelet modulation technique has been developed and modeled for the 1-ph 4-pulse H-bridge voltage source DC-AC inverter. Simulation has been performed to analyze its performance for R and RL Load. Simulation results have been compared with square pulse and PWM switching techniques in the identical load conditions, the current harmonics are more compared to PWM. A different kind of derived Scaling function is proposed to be used in future which can give a much more improved result.

References

[1] Dr. P. S. Bimbhra, "Power Electronics," 4th ed. Khanna Publishers, Delhi:2009

[2] Jai P. Agrawal, " Power electronic systems – Theory and Design," 1st ed. Pearson Education Asia, Singapore:2001

[3] S. A. Saleh, C. R. Moloney and M. A. Rahman, "Experimental Performance of the Single-Phase Wavelet-Modulated Inverter," IEEE Trans. Power Electronics, vol. 26, no. 9, pp. 2650 – 2600, Sep. 2011.

[4] http://en.wikipedia.org/wiki/Wavelet

[5] S. A. Saleh, C. R. Moloney and M. A. Rahman, "Developing a Non – Dyadaic MRAS for Switching DC – AC inverters," in Proc. IEEE 12th DSP Conf. , Jackson Lake Lodge, WY, Sep. 2006, pp. 544 – 549.

[6] S. A. Saleh, C. R. Moloney and M. A. Rahman, "Development and Testing of Wavelet Modulation for Single-Phase Inverters," IEEE Trans. Ind. Electronics, vol. 56, no. 7, pp. 2588 – 2599, July 2009.

[7] S. A. Saleh, "The Implementation and Performance Evaluation of 3 VS Wavelet Modulated AC – DC Converters," IEEE Trans. Power Electronics, vol. 28, no. 3, pp. 1096 – 1106, March 2013.

[8] Stephane Mallat, "a Wavelet Tour of Signal Processing," 2nd ed. Academic Press, London: 1999

[9] Agostino Abbate, Casimer M. DeCusatis and Pankaj K. Das, " Wavelet and Subbands Fundamental and Applications," Birkhauser, Boston:2002

[10] http://en.wikipedia.org/wiki/Multiresolution_analysis

[11] Stephen J. Chapman, "MATLAB Programming for Engineers," 3rd ed. Thomson Learning, Singapore: 2005.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

Adiabatic Vedic Multipler Design Using Chinese Abacus Approach

Diksha Ruhela[1], Dr. Malti Bansal[2]

M.Tech (VLSI), Electronics & Communication Engineering, Delhi Technological University, Delhi, India[1]

Assistant Professor, Electronics & Communication Engineering, Delhi Technological University, Delhi, India[2]

Abstract: In this paper, we discuss the implementation of Vedic multiplier with Chinese Abacus adder design, using Reversible Logic Gates. The power consumption of the Vedic multiplier is low as it generates all partial products and their sum in one step. Since high radix of Chinese Abacus adder reduces the carry propagation delay, it is observed here that the proposed design increases the speed of operation manifold. The proposed work is implemented on the Xilinx FPGA device, Spartan-3E. The results show that multiplier implemented using Chinese Abacus approach is quite efficient in terms of area, time and speed

Keywords: Vedic Multiplier, Chinese Abacus Adder, propagation delay, Reversible Logic.

I. INTRODUCTION

Multiplier is one of the important blocks in almost all the arithmetic logic units. These multipliers are mostly used in the fields of Digital Signal Processing (DSP), Fast Fourier Transforms, convolution, filtering and microprocessor applications. A system's performance is generally determined by the performance of the multiplier, because the multiplier is generally the slowest element in the system. Furthermore, it is generally the most area consuming. Hence, optimizing the speed and area of the multiplier is a major design issue. A high speed and area efficient multiplier can be achieved by using Vedic mathematics. In this work, we have implemented the Vedic multiplier using Chinese Abacus Adder with and without using Reversible logic gates.

Reversible logic is one of the promising fields for future low power design technologies. Since one of the requirements of all DSP processors and other embedded devices is to minimize power dissipation, multipliers with high speed and lower dissipation are critical. This paper proposes an implementation of Reversible *Urdhva Tiryakbhayam* (UT) Multiplier which consists of two cardinal features. One is the fast multiplication feature derived from Vedic algorithm *Urdhva Tiryakbhayam* and another is the reduced heat dissipation by the virtue of implementing the circuit using reversible logic gates. The paper is partitioned into six sections. Section II gives introduction on reversible logic. Section III explains the *Urdhva Tiryakbhayam* (UT) algorithm. Section IV explains the Chinese Abacus Adder algorithm, Section V elaborates on the design aspects of Reversible UT Multiplier. Section VI gives the simulation results. Conc1usions and references follow

II. REVERSIBLE LOGIC

Reversible logic is a promising computing design paradigm which presents a method for constructing computers that produce no heat dissipation. Reversible computing emerged as a result of the application of quantum mechanics principles towards the development of a universal computing machine. Specifically, the fundamentals of reversible computing are based on the relationship between entropy, heat transfer between molecules in a system, the probability of a quantum particle occupying a particular state at any given time, and the quantum electrodynamics between electrons when they are in proximity. The basic principle of reversible computing is that an objective device with an identical number of input and output lines will produce a computing environment where the electrodynamics of the system allow for prediction of all future states based on known past states, and the system reaches every possible state, resulting in no heat dissipation. A reversible logic gate is an N-input, N-output logic device that provides one to one mapping between the input and the output. It not only helps us to determine the outputs from the inputs but also helps us to uniquely recover the inputs from the outputs.

The following are the important design constraints for reversible logic circuits:

1. Reversible logic gates do not allow fan-outs.
2. Reversible logic circuits should have minimum quantum cost.
3. The design can be optimized so as to produce minimum number of garbage outputs.
4. The reversible logic circuits must use minimum number of constant inputs.
5. The reversible logic circuits must use a minimum logic depth or gate levels.

The basic reversible logic gates encountered during the design are listed below:

1. Feynman Gate [1]:

It is a 2x2 gate and its logic circuit is as shown in the figure. It is also known as Controlled Not (CNOT) Gate. It has quantum cost one and is generally used for Fan Out purposes.

2. Peres Gate [2]:

It is a 3x3 gate and its logic circuit is as shown in the figure. It has quantum cost four. It is used to realize various Boolean functions such as AND, XOR.

Copyright to IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

3. Fredkin Gate [3]:

It is a 3x3 gate and its logic circuit is as shown in the figure. It has quantum cost five. It can be used to implement a Multiplexer.

4. HNG Gate[4]:

It is a 4x4 gate and its logic circuit is as shown in the figure. It has quantum cost six. It is used for designing ripple carry adders. It can produce both sum and carry in a single gate thus
minimizing the garbage and gate counts.

5. NFT Gate[5]:

The 3*3 New Fault Tolerant gate (NFT) with quantum cost of 5 has worst case delay of 3 and it has better correction capability. The output states map to the inputs in this manner.

A = A XOR B,
B = AC' XOR B'C, and
C = AC' XORB C

6. F2G Gate[5]

The 3*3 Feynman double gate gate with quantum cost of 2 has worst case delay of 3 and it has better correction capability

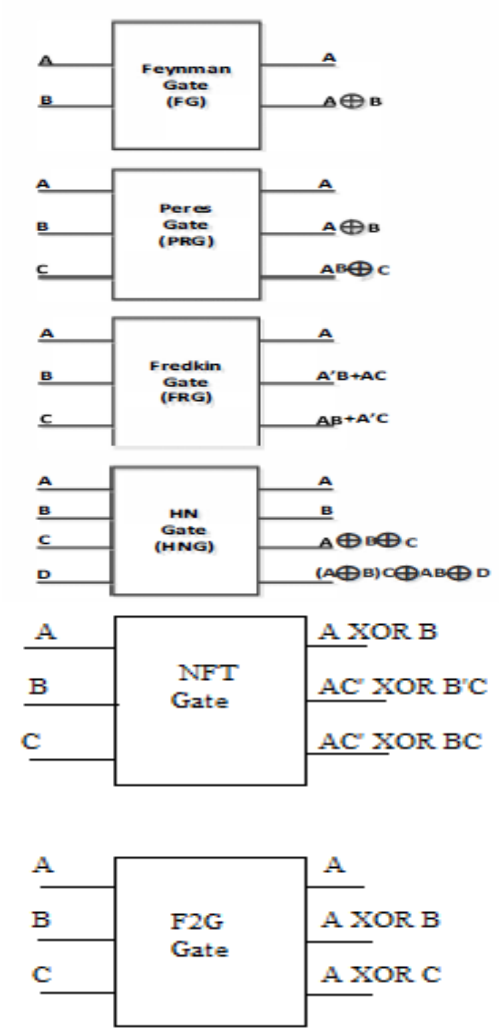


Figure1: Reversible Logic Gates[1][2][3][4][5]

III. URDHAVA – TIRYAGBHYAM

The multiplier is based on an algorithm *Urdhva Tiryakbhyam* (Vertical & Crosswise) of ancient Indian Vedic Mathematics. *Urdhva Tiryakbhyam* Sutra is a general multiplication formula applicable to all cases of multiplication. It literally means "Vertically and crosswise". It is based on a novel concept through which the generation of all partial products can be done with the concurrent addition of these partial products.

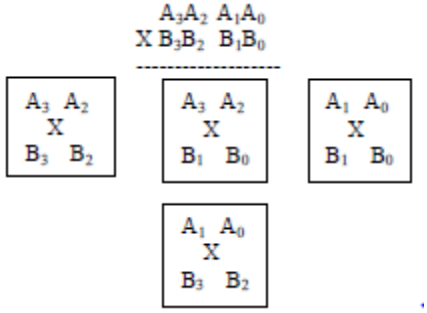


Figure 2: Block diagram of 4x4 vedic multiplication.

For NXN multiplication unit, we require four N/2 bit multipliers and three N bit full adders. 4x4 bit multiplication has been shown in Fig.3. High speed of multiplier depends mainly upon speed of the adder units used.

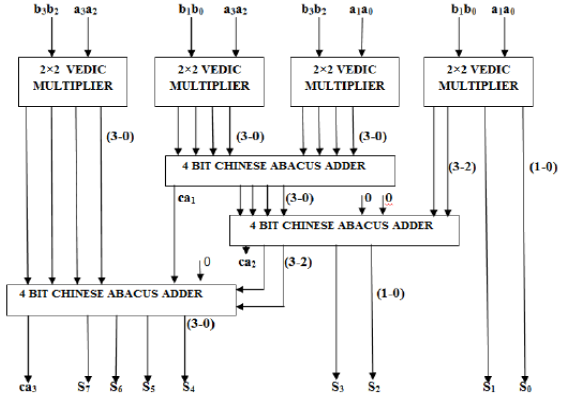


Figure3: Block diagram of proposed 4x4 Vedic multiplier using Chinese Abacus Adder

IV. CHINESE ABACUS

Chinese Abacus is the ancient approach used for performing mathematical calculation in many parts of the world, especially in China. It is still used by many small enterprises for their day to day work. Due to its high speed of operation, it give high competition to electronic pocket calculator. Similar efficiency can be attained in electronic version.

The first digital arithmetic circuit employing the Chinese abacus approach were proposed by Maloberti et al. [6-8]. The architecture of radix-4 abacus adder is presented in [6]. The architecture of radix-2 abacus adder is presented in [7] and [8]. The most severe drawback of Maloberti's adder in radix-4 design is the serial addition of each bead by the Shift-Up module. The serial addition of each bead by the Shift-Up module can be improved by the parallel addition

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

module of the proposed architecture. The serial addition of each bead by the Shift-Up module will moreover induce more than nine pass transistor chains, due to its counting method. The delay time of the signal waveform becomes long.

Radix-4 adder helps in reducing the excessive delay by reducing the number of carries [9]. A basic column element of the Chinese abacus is depicted in figure 4(a). The configuration showed in figure 4(a) represents the number six. Each column element has one higher bead with a weight, of five and four lower beads with a weight of one. The key feature of the Chinese abacus is the use of one bead with weight five. This allows the operator to minimize the transmission of rests.

In this paper, we use abacus adder in each column having three beads higher and three beads lower, each having weight of four and one each, respectively. The basic element is able to represent decimal number in the range 0 to 15. Figure 4(b) depicts the abacus representation of decimal number 9. In this methodology, the 4-bit adder only contributes two ripple carries. One is internal carry from lower beads to higher beads, the other is external carry from this column element to a next neighbour column element.

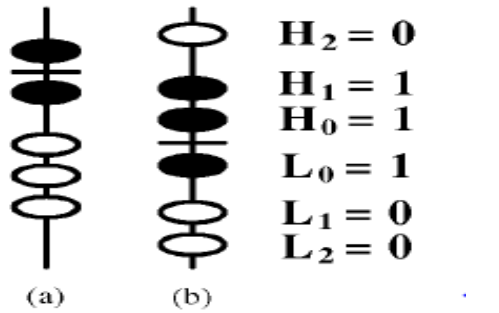


Figure 4. (a) Chinese abacus coding with base 10 of the decimal number 6, (b) the proposed Chinese abacus adder coding with base 16 of the decimal number 9 [7].

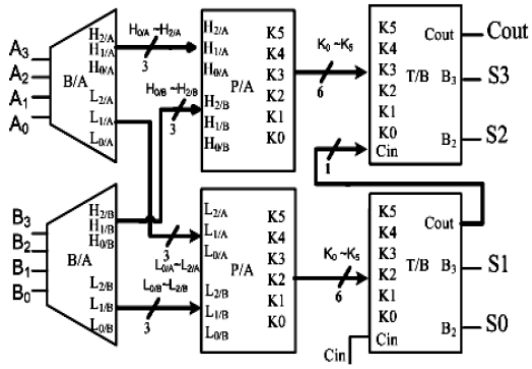


Figure 5. Block diagram of the 4-bit abacus adder[10].

Block diagram of radix-4 abacus adder is depicted in figure 5. As shown in figure 5, the function of radix-4 abacus adder is completed in four phases. In first phase binary input is converted into abacus output known as B/A(binary to abacus) module. In second phase, parallel additon takes place known as P/A(parallel addition module). In the last (third) phase thermometric input converted into binary output known as T/B(thermometric to binary) transformation module.

V. IMPLEMENTATION OF PROPOSED DESIGN AND SIMULATION RESULTS

The 4x4 Vedic multiplier is design using the Chinese abacus methodology. It may simply use three 4-bit abacus adders, shown in figure 3. We design all component of multiplier using reversible logic gates.

The digital logic implementation of the 2X2 Urdhva Tiryakbhayam multiplier using the conventional logic gates [11] is as shown in figure 6. The expressions for the four output bits are given below figure 6. The block diagram of reversible logic gate implementation of 2x2 UT multiplier is as shown in figure 7. This design does not consider the fanouts.

The circuit requires a total of six reversible logic gates out of which five are Peres gates and remaining one is the Feynman Gate. The quantum cost of the 2X2 Urdhva Tiryakbhayam Multiplier is enumerated to be 21. The number of garbage outputs is 9 and number of constant inputs is 4.

The Reversible 4X4 Urdhva Tiryakbhayam Multiplier design emanates from the 2X2 multiplier. The block diagram of the 4X4 Vedic Multiplier is presented in the figure 3. It consists of four 2X2 multipliers, each of which processes four bits as inputs; two bits from the multiplicand and two bits from the multiplier. The lower two bits of the output of the first 2X2 multiplier are entrapped as the lowest two bits of the final result of multiplication. Two zeros are concatenated with the upper two bits and given as input to the four bit abacus adder.

The other four input bits for the abacus adder are obtained from the second 2X2 multiplier. Likewise, the outputs of the third and the terminal 2X2 multipliers are given as inputs to the second four bit abacus adder. The outputs of these four bit adders are in turn 5 bits each, which need to be summed up. This is done by a five bit abacus adder which generates a six bit output. These six bits form the upper bits of the final result.

The Chinese abacus adder is consummated (realized) using the NFT and F2G Gate. Single NFT Full Adder (SNFA**)** is a Fault Tolerant full adder circuit which consists of one New Fault Tolerant (NFT) gate and three Feynman Double (F2G) gates where the quantum cost is 11 and the total number of garbage outputs is 3.

This design also does not take into consideration the fanouts of the gates. For this design the quantum cost is computed to be 162, the total number of gates used will be 36, the number of garbage outputs will be 62 and the number of constant inputs will be 29.

Copyright to IJARCCE
DOI 10.17148/IJARCCE.2015.4418

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

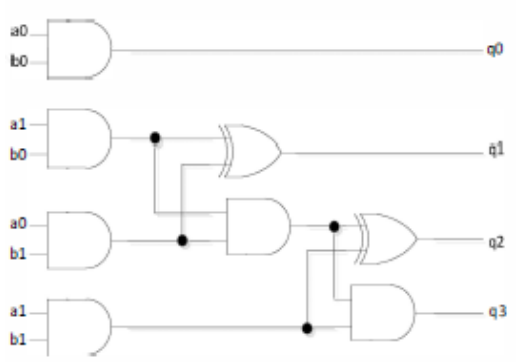


Figure 6: Conventinal Logic Implementation of 2x2 UT multiplier

q0 = a0.b0
ql= (a1.b0) xor (a0.bl)
q2= (a0.al.b0.bl) xor (al.bl)
q3= a0.al.b0.bl

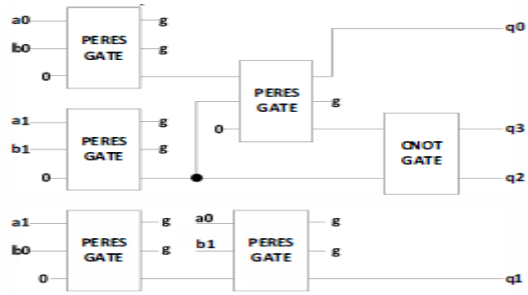


Figure 7: Reversible logic gate implementation of 2x2 UT multiplier.

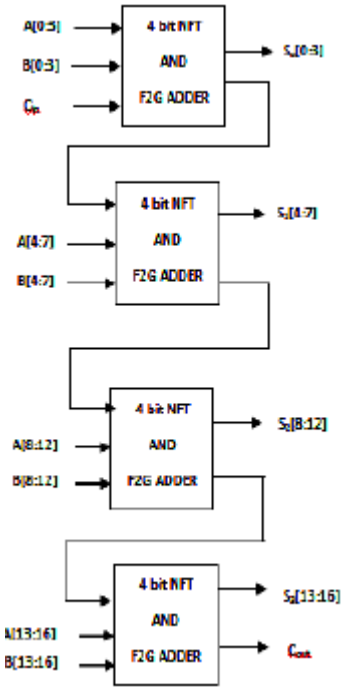


Figure 8: Reversible logic gate implementation of Chinese Abacus Adder.

Simulation of the proposed work is done using Modelsim SE 6.5. Synthesis and Implementation is done using Xilinx Spartan-3E FPGA Board of device family xc3s50-4-tq144. The subsequent figures show the simulation, RTL and Technical Schematic of the multiplier. Results show that 4x4 UT vedic multiplier designed in this work using reversible logic gates, consumes less area and has high speed as compared to 4x4 UT vedic multiplier designed using conventional logic gates. Results also show that 4x4 UT vedic multiplier designed using reversible logic gates is 23% faster and takes less area than 4x4 UT vedic multiplier design without using reversible logic gates

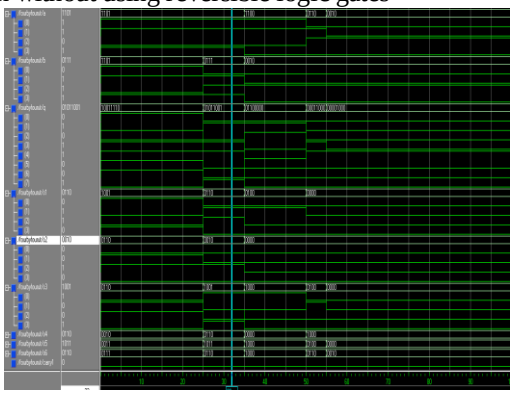


Figure 9: Simulation result for 4x4 bit Vedic multiplier designed using reversible Logic gates

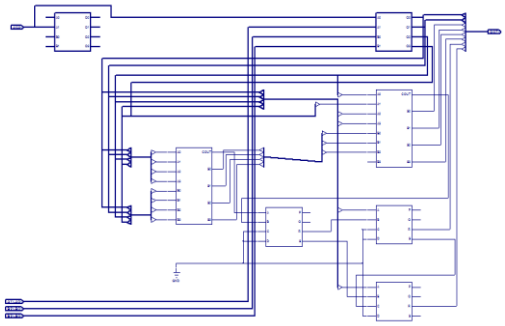


Figure 10: RTL view of 4x4 bit UT Vedic multiplier designed using reversible logic gates.

TABLE I: SIMULTAION RESULTS OF DEVICE UTILIZATION

Methods	**UT Vedic multiplier without reversible logic gates**	**UT Vedic multiplier with reversible logic gates**
Number of slices	36 out of 768 (4%)	28 out of 768 (3%)
Number of 4 input LUTs	65 out of 1536 (4%)	49 out 1536 (3%)
Number of bonded IOBs	17 out of 97 (17%)	16 out of 97 (17%)
Net propagation delay	20.705 ns	15.830 ns
Frequency	48.297MHz	63.17MHz

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

VI. CONCLUSION

The use of Vedic multiplier approach results in a competitive technique as compared respect to conventional fast multiplier. The simulation results show that this approach of Vedic multiplier using Chinese Abacus adder is very efficient for low-power, high-speed applications. This architecture is also easy for pipeline implementation. Another advantage is that the methodology may reduce the number of ripple carries and partial product generation in many steps. It can be inferred that Vedic multiplier using Chinese Abacus adder is quite efficient as compared to the conventional multiplier, for multiplication operation.

REFERENCES

[1]. R. Feynman, "Quantum Mechanical Computers," Optics News, Vol.1l, pp. 11-20, 1985.
[2]. Peres, "Reversible logic and quantum computers", Phys. Rev. A 32 (1985) 3266-3276.
[3]. E. Fredkin and T. Toffoli,"Conservative Logic", Int'l 1 Theoretical Physics Vo121, pp.219-253, 1982.
[4]. Shams, M., M. Haghparast and K. Navi, Novel reversible multiplier circuits in nanotechnology. World Appl. Sci. J., 3(5): 806-810.
[5]. Shailja Shukla, Tarun Verma and Rita Jain, "Design of 16 Bit Carry Look Ahead Adder Using Reversible Logic", International Journal of Electrical, Electronics and Computer Engineering 3(1): 83-89(2014).
[6]. Franco Maloberti and Chen Gang, "The Chinese Abacus method: can we use it for digital arithmetic", Proceedings of the 8th Great Lakes Symposium on VLSI, pp. 192 - 195 19-21 Feb. 1998.
[7]. Franco Maloberti and Chen Gang, "Use of the Chinese Abacus method for digital arithmetic functions", Proceedings of the 1998 IEEE International Symposium on Circuits and Systems, Vol. 5, pp. 213 - 216, vol.5, 31 May - 3 June 1998.
[8]. Franco Maloberti and Chen Gang, "Performing Arithmetic Functions with the Chinese Abacus Approach," IEEE Transactions on circuits and systems-II: Analog and digital signal processing, vol. 46, no. 12, pp. 1512- 1515, Dec. 1999.
[9]. Andreev, B.D., Titlebaum, E., Friedman, E.G., "Tapered Transmission Gate Chains for Improved Carry Propagation,"MWSCAS-2002. The 45th Midwest Symposium on Circuits and Systems, Volume 3, pp. 449 - 452, Aug. 4-7, 2002.
[10]. Behrooz Parhami "Computer Arithmetic-Algorithms and Hardware Designs", Oxford University Press, Inc (2000).
[11]. G Ganesh Kumar and V Charishma, "Design of high speed vedic multiplier using vedic mathematics techniques", ltn'l J. of Scientific and Research Publications, Vol. 2 Issue 3 March 2012.
[12]. Maloberti, F., Chen Gang, "Functions with the Chinese Abacus Approach", IEEE Transaction on circuits and systems-II: Analog and digital signal processing, vol. 46, no. 12, pp. 1512 – 1515, Dec.1999.
[13]. Shu-Chung Yi, Kun-Tse Lee, Jin-Jia Chen, Chien-Hung Lin, Chuen-Ching Wang, "The new architecture of radix-4 Chinese abacus adder", Proceedings of the 36th International Symposium on Multiple-Valued Logic (ISMVL'06) 2006 IEEE.
[14]. Shu-Chung Yi, et al.,"Comments on Performing Arithmetic Functions with the Chinese Abacus Approach", IEEE Transaction on circuits and systems-II.
[15]. Rakshith Saligram and Rakshith T.R. "Design of Reversible Multipliers for linear filtering Applications in DSP", International Journal of VLSI Design and Communication systems, Dec-2012.
[16]. .Harpreet Singh Dhillon Abhijit Mitra, "A Digital Multiplier Architecture using Urdhava Tiryabhyam Sutra of Vedic Mathematics", International Conference on Design & Technology of Integrated Systems in Nanoscale Era, 2008.
[17]. Dr. Malti Bansal, Diksha Ruhela, "High Speed & Area Efficient Vedic Multiplier using Adiabatic Logic", Journal of Basic and Applied Engineering Research Volume 1, Number 11; October-December 2014 pp. 14-17.

Orig Life Evol Biosph
DOI 10.1007/s11084-015-9437-2

PREBIOTIC CHEMISTRY

Affinity of Smectite and Divalent Metal Ions (Mg^{2+}, Ca^{2+}, Cu^{2+}) with L-leucine: An Experimental and Theoretical Approach Relevant to Astrobiology

Pramod Pandey[1] · Chandra Kala Pant[1] · Kavita Gururani[1] · Priyanka Arora[2] · Neetu Pandey[2] · Preeti Bhatt[1] · Yogesh Sharma[1] · Jagmohan Singh Negi[3] · Mohan Singh Mehata[4]

Received: 3 December 2014 / Accepted: 17 March 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Earth is the only known planet bestowed with life. Several attempts have been made to explore the pathways of the origin of life on planet Earth. The search for the chemistry which gave rise to life has given answers related to the formation of biomonomers, and their adsorption on solid surfaces has gained much attention for the catalysis and stabilization processes related to the abiotic chemical evolution of the complex molecules of life. In this communication, surface interactions of L-leucine (Leu) on smectite (SMT) group of clay (viz. bentonite and montmorillonite) and their divalent metal ion (Mg^{2+}, Ca^{2+} and Cu^{2+}) incorporated on SMT has been studied to find the optimal conditions of time, pH, and concentration at ambient temperature (298 K). The progress of adsorption was followed spectrophotometrically and further characterized by FTIR, SEM/EDS and XRD. Leu, a neutral/non polar amino acid, was found to have more affinity in its zwitterionic form towards Cu^{2+}- exchanged SMT and minimal affinity for Mg^{2+}- exchanged SMT. The vibrational frequency shifts of —NH_3^+ and —COO^- favor Van der Waal's forces

Electronic supplementary material The online version of this article (doi:10.1007/s11084-015-9437-2) contains supplementary material, which is available to authorized users.

✉ Pramod Pandey
pp.chem24@gmail.com

✉ Chandra Kala Pant
ckp.astrochem@gmail.com

[1] Chemical Laboratory, Department of Chemistry, DSB Campus, Kumaun University, Nainital 263002 Uttarakhand, India

[2] Photophysics Laboratory, Department of Physics, DSB Campus, Kumaun University, Nainital 263002 Uttarakhand, India

[3] Herbal Analytical Laboratory, Herbal Research and Development Institute, Mandal, Gopeshwar 246401 Uttarakhand, India

[4] Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi 110042, India

Published online: 08 May 2015

during the course of surface interaction. Quantum calculations using density functional theory (DFT) have been applied to investigate the absolute value of metal ion affinities of Leu (Leu—M^{2+} complex, M=Mg^{2+}, Ca^{2+}, Cu^{2+}) with the help of their physico-chemical parameters. The hydration effect on the relative stability and geometry of the individual species of Leu—$M^{2+}\times(H_2O)_n$, (n =2 and 4) has also been evaluated within the supermolecule approach. Evidence gathered from investigations of surface interactions, divalent metal ions affinities and hydration effects with biomolecules may be important for better understanding of chemical evolution, the stabilization of biomolecules on solid surfaces and biomolecular-metal interactions. These results may have implications for understanding the origin of life and the preservation of biomarkers.

Keywords Amino acids · Smectite · Metal ions Surface interaction · DFT Origin of life

Introduction

The origin of life is generally presumed to require a local concentration of life-forming biomolecules (e.g. amino acids, nucleic acid bases, phosphate and sugars etc.), which were either supplied abundantly or sufficiently stable to accumulate under primitive Earth conditions. Prebiotic syntheses of biomolecules from chemical precursors often suffer from low yields (Lucrezia et al. 2007; Orgel 2004; Lambert 2008; Shapiro 1999), thus the steady-state concentrations of amino acids (AAs) in the primitive oceans has been estimated to be on the order of 4 to 10^{-7} mM (Lahav and Chang 1976; Stribling and Miller 1987).The stability of these molecules under the conditions (viz. ultra-violet, thermal, electric discharge etc.) prevalent on Earth when life started ~4.4–3.5 Ga ago may also have been an impediment to their accumulation (Cleaves et al. 2012).

Bernal (1951) suggested a possible solution to these problems and postulated concentration of biomolecules on mineral surfaces like metal oxides. Metal oxides and silicates are widely distributed on the Earth's surface and also reported on another planet like Mars (Stephenson et al. 2013; Carr and Head 2010). Bentonite (BNT) and montmorillonite (MMT) clays, well known for their cation exchange capacities, had been widely investigated at different conditions of pH, temperature and concentrations to explore Bernal's ideas regarding their potential roles as concentrating surfaces and catalysts for polymerization (Zaia 2012; Lambert 2008 and references therein it). The sorption behavior of SMT surfaces is primarily governed by variables such as the interlayer conditions, presence of cations, water, etc. (Prabhakar et al. 2007).

Metal ions are known for their affinity to form peptide bonds in relevance to the prebiotic evolution of biomolecules (Rode et al. 1993; Pant et al 2009; Gururani et al. 2012a; Remko and Rode 2000). Metal ion-exchanged SMTs are good adsorbents for amino acids, nucleic acid bases, sugars etc., and also catalyst for biomolecule formation (Lambert 2008; Kalra et al. 2003; Gururani et al. 2012b). A detailed mechanism behind prebiotic mineral-organic interfacial processes and its potential implications for the origins of life have been reviewed by Cleaves et al. (2012).

Recently, the role of metal ions in chemical evolution has been bolstered with advances in computational chemistry (Remko and Rode 2000; Rode et al. 1993: Merz et al. 2014; Cleaves et al. 2012). The possible interaction of metal ions with biomolecules is often investigated with the help of high-pressure mass spectrometry along with guided ion beam mass spectrometry

(Kebarle and Hogg 1965; Rodgers and Armentrout 2000). However, quantum mechanical calculations give better insight into their possible complex formation mechanism, solvation and counter ions effects of the different complexes (Marino et al. 2001). Hydrated biomolecular complexes serve as simpler models for more complex protein–metal systems in an aqueous environment (Remko et al. 2011).

Among other prebiotically formed AAs, Leu has been reported from exogenous (Pizzarello 2009) as well as endogenous (Zaia 2012) sources. However, to the best of our knowledge, a systemic experimental study of the surface interactions of Leu on Ca^{2+}-, Mg^{2+}- and Cu^{2+}-exchanged SMTs (BNT and MMT), separately, has not been realized yet. In this communication, we investigate the adsorption behavior of different metal ion exchanged SMTs because such studies related to the biomolecular interaction on prebiotically available solid surface and their affinity towards silicates, are relevant to astrobiology (Lambert 2008; Cleaves et al. 2012). The stability of the adsorbate under different experimental conditions is also a major concern for astrobiological investigation as silicates are also known for their catalytic role in polymerization and decomposition of biomolecules (Zaia 2012; Marshall-Bowman et al. 2010; Ferris et al. 1996). Thus, the stability of Leu was also investigated under different sorption conditions.

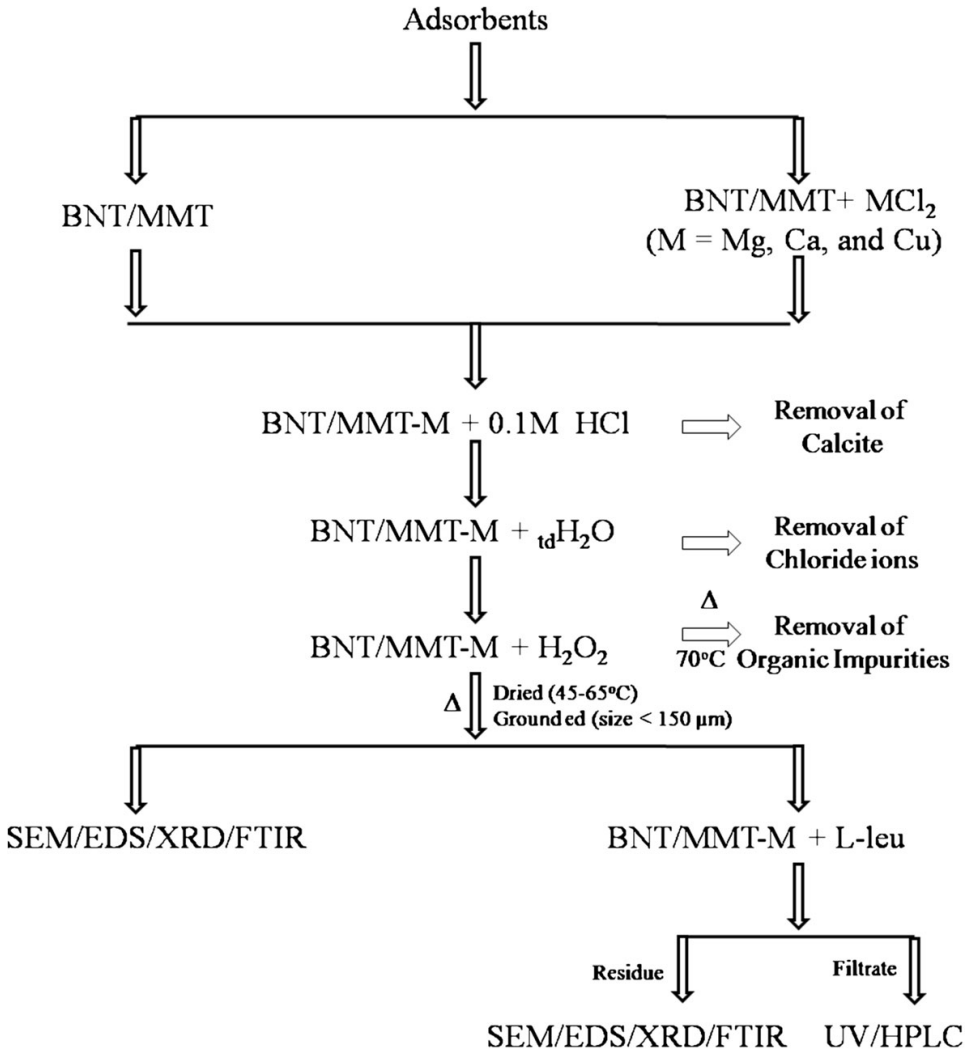
Computational studies of metal-amino acid complexes (Leu—M^{2+}) were conducted using Gaussian 03 (Frisch et al. 2003) to understand the affinity of metal ions with biomolecules and the effect of hydration on Leu—$M^{2+}\times(H_2O)_n$, (n =2 and 4) complex formation, which has not been studied yet theoretically. The details of local interactions between metal ions and AAs are of prime importance in living systems (Marino et al. 2001).

Experimental

All reagents were of analytical grade. BNT and MNT K10 clays were purchased from E. Merck, India and Sigma Aldrich, Germany respectively, while L-leucine (Leu), H_2O_2, $MgCl_2$, $CuCl_2$ were from Qualigens, India and $CaCl_2$ was obtained from CDH India. Triple distilled-deionized water (pH 6.8±0.1) was used throughout the study. All containers, measuring equipment and assemblies were of borosilicate glass. A JascoV-550, UV/VIS/NIR spectrophotometer was used for the absorbance measurements. The pH was recorded on Hanna pH meter (model pHep, accuracy: ±0.1 pH unit).

The sequential steps involved in the preparation/purification of adsorbents during analysis are shown in scheme I, and described by Prabhakar et al. 2007 and Kalra et al. 2003. Ca^{2+}, Cu^{2+} and Mg^{2+} exchanged BNT/MMT were prepared by the saturation method as mentioned earlier (Kalra et al. 2003). All the adsorbents (Scheme I) were ground, sieved using a 100 mesh size sieve (SETHI BSS 100) and then dried in vacuum desiccators. The color of the respective SMTs have been found to be 10R8/1 (white) and 7.5YR7/2 (light brown) on Munsell Notations, for MMT and BNT respectively.

The purity of adsorbents (BNT/MMT and their divalent cation-exchanged forms) were determined by scanning electron microscopy (SEM) and energy-dispersive X-ray spectrometry (EDS) using a Hitachi S-3700 N Instrument using an acceleration voltage 15.0 kV with a take-off angle 44.1°,and a ZEISS EHT 20 kV instrument. Purity was further checked by X-ray diffraction (relative intensity and inter-planar spacing) data, which were in good agreement with reported values (Caglar et al. 2009). X-ray powder diffraction analysis of adsorbent and adsorption adducts were performed on a Thermal X-TRA ARL diffractometer equipped with a



Scheme I Graphical representation of different steps involved in the preparation/purification of adsorbents and different analysis

'Peltier' detector, using Cu Kα radiation (wavelength=1.5406 Å) operated at 30 mA and 40 kV and graphite monochromator detection between range 0-40° 2θ angle. Vibrational studies were carried out using a PerkinElmer DX-II FT-IR spectrophotometer equipped with DTGS detector. The IR spectra were collected at 4 cm^{-1} resolution and 100 scan with frequency range from 400 to 4000 cm^{-1}.

The adsorption of Leu on different adsorbents in aqueous medium was studied as function of pH (4.0–9.2), time (1 min – 72 h) and concentration (1 mM to 10 mM) of adsorbate. Adsorption of Leu in varying concentration over different pH was studied in order to obtain saturation point by adding acetate-acetic acid (pH 4.0–6.0) or borax-boric acid (pH 6.6–9.2) buffer to the AA solution containing adsorbents, keeping in mind that the buffer solution should be a very poor ligand so that stable complex formation with the mineral could be avoided and this was confirmed by using different buffer concentration on Leu solution.

Buffer solution of Leu (5 ml) at different pH values was added to the adsorbents (50 mg) in separate conical flasks (10 ml) under air, stirred mechanically using magnetic stirrer and allowed to stand at room temperature (298 K) for different time intervals to determine the optimum conditions for adsorption. The solutions were then centrifuged at 2000 rpm for

25 min. The supernatant was decanted leaving the SMT as a residue. The supernatant in each experiment was used for quantitative estimation of AA content after adsorption and the SMT residue was separated, washed (3 times) and then, dried at ambient temperature under vacuum. Further, it was ground/ sieved and used for SEM/EDS, FTIR and XRD studies (Scheme I). The concentration of adsorbate before and after adsorption was determined using UV difference spectra (at 201 nm) in quartz cuvettes with a 1 cm path length using $_{td}H_2O$ as reference. Same amount of respective buffer had also been added into the $_{td}H_2O$ which was used as reference.

The amount of AA adsorbed under different conditions of pH, time and concentration was calculated from the difference between the initial AA concentration and the concentration after adsorption in each case. The initial concentration of AA and the quantity adsorbed were used to obtain the adsorption isotherms (Pandey et al. 2013). Percentage binding was calculated as follows-

$$\%\,binding = \frac{C_e - C_f}{C_e} \times 100 \tag{1}$$

Where, C_e is the initial concentration of adsorbate and C_f is the final concentration after adsorption. The theoretical limit of the concentration of adsorbate (Leu) available for adsorption is given (Fig. 1a).

Chromatographic techniques (Paper chromatography and HPLC) were used to measure the degradation products (if any) of Leu after adsorption studies as described in our previous work (Pandey et al. 2013).

Computational Study

Density functional theory (DFT) was used to investigate the interaction of Leu with metal cations (Ca^{2+}, Mg^{2+} and Cu^{2+}). Becke's three parameter exchange functional with a Lee, Yang and Parr type nonlocal correlation functional hybrid Hartee-Fock DFT method (B3LYP),was applied for metal ions and Leu with the all-electron 6-31G+(dp) internal basic set in Gaussian 03w (Becke 1993; Lee et al. 1988; Hohenberg and Kohn 1964). All the geometries were

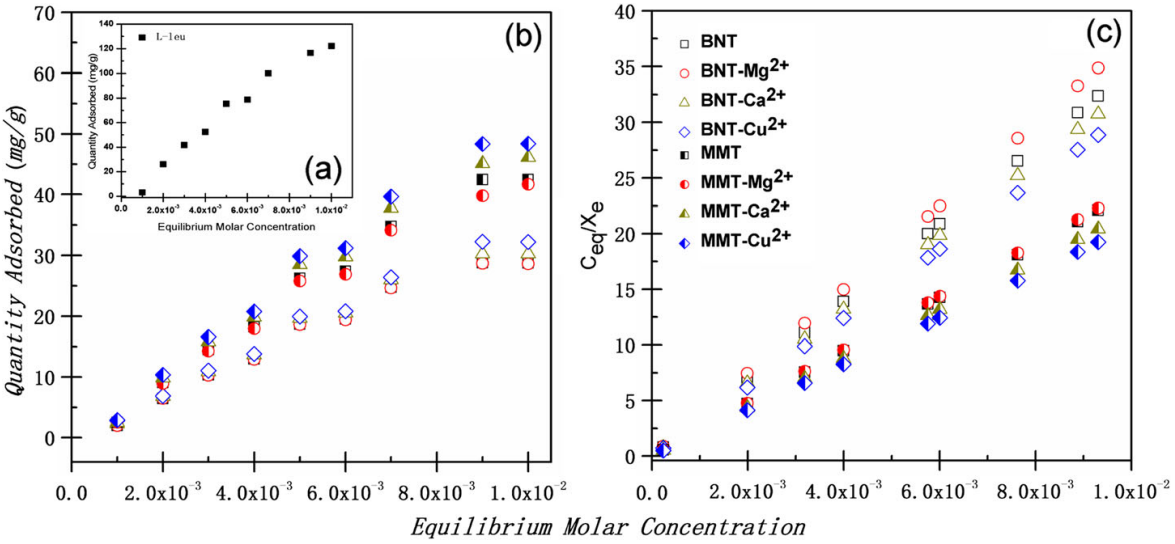


Fig. 1 Graph showing, (**a**) theoretical limit of availability of Leu for complete adsorption by 50 mg of adsorbents, (**b**) adsorption behaviour as a function of quantity adsorbed versus equilibrium molar concentration, and (**c**) Langmuir adsorption isotherm for different adsorbents for equilibrium time 8 h and pH 6.0

optimized and frequency calculations were used to cross-check the global minima of all the chemical species under study, where no imaginary frequencies were observed. Thermodynamic parameters are also obtained through frequency calculation. Open shell calculations using a spin-unrestricted formalism have been used for Leu—Cu^{2+}complex as Cu^{2+} is an open shell system with a $d^9(^2D)$ ground state (Remko et al. 2011). All the computational analysis were carried out with Intel(R) Core (TM) i5-3230 M CPU@2.60GHz PC on Windows-7 operating system. The formation of Leu—$M^{2+}\times(H_2O)_n$, (n=2 and 4) complexes can be described following, which has been used to compute the metal ion affinity as a negative enthalpy variation (ΔH) for the reaction-

$$\begin{array}{l} Leu\ +\ M^{2+} + (H_2O)_n \rightarrow Leu—M^{2+} \times (H_2O) \\ M = Mg^{2+},\ Ca^{2+} and\ Cu^{2+},\ n = 0,\ 2\ and\ 4. \end{array} \quad (2)$$

Results and Discussion

The adsorption of Leu on MMT/BNT and cation exchanged MMT/BNT, in aqueous medium was studied as a function of pH (4.0–9.2), temperature (298 K) and concentration (1 mM – 10 mM) of the adsorbate in order to find out the conditions of maximum adsorption. Moreover, the main aim is to study the role of divalent cation (Cu^{2+}, Mg^{2+} and Ca^{2+}) exchanged SMT as an adsorbent under prebiotic conditions.

The amount of adsorbate was determined from the standard concentration curve obtained from the optical density of the different Leu solutions monitored at their respective peak (λ_{max}) of absorption spectrum. Linear behaviour (r^2=0.97) was observed between the absorbance and concentration of Leu (Supplementary information S.1). Moreover, Leu adsorption was confirmed using EDS which showed the presence of C and N in addition to the elements present in the SMT.

The optimum conditions of surface interaction with respect to time and pH showed that the amount adsorbed depends on the contact time, and pH, where adsorption equilibrium is attained within 8 h (Supplementary information S.2). The optimum pH was found to be 6.0 which is also, isoelectric point of Leu (Supplementary information S.3).

The theoretical limit of the availability of total adsorbate (Leu) in 5 ml for 100 % adsorption by the 50 mg of SMT is shown in Fig. 1a. The adsorption isotherms (plot between equilibrium concentrations (C_{eq}) versus quantity adsorbed) are provided in Fig. 1b. The initial portion of the isotherm (From 1 to 6 mM of Leu) represents a linear relationship between the quantity adsorbed and C_{eq} of adsorbate, while at higher concentration (6 mM—10 mM), the isotherm saturates and adsorption becomes independent of Leu concentration. Langmuir type adsorption is observed (Fig. 1c) assuming the formation of a monolayer of adsorbate molecules on the surface of the adsorbent. The Langmuir parameters were determined form the equation below (Pandey et al. 2013).

$$\frac{C_{eq}}{X_e} = \frac{1}{K_L.X_m} + \frac{C_{eq}}{X_m} \quad (3)$$

Where C_{eq} is the equilibrium concentration of Leu.K_L is a constant which is a function of the adsorption energy *i.e.* the enthalpy of adsorption coefficient ($K_L=Ae^{-\Delta H/RT}$); X_e is amount of amino acid (mg) adsorbed per gram of SMT;X_m is the amount of Leu required for a definite weight of SMT for complete surface coverage.

The adsorption parameters (X_m and K_L) were calculated from the slope and intercept obtained from the graph of C_{eq}/X_e versus C_{eq} (Fig. 1c and Eq. 3). The goodness of linear fit of the Langmuir adsorption isotherm (Fig. 1c) was judged by the linear regression coefficients (r^2) which were ~0.99 (Table 1). Adsorption data were collected in triplicate and ~3.0–7.8 % error was observed in this study. Mean values are used to obtain isotherm (Fig. 1b, c).

Percentage binding of Leu on the two SMT is shown in Table 1, using eq. 1. Leu is ~24 % adsorbed on BNT while ~34 % is adsorbed for MMT, and same trend of the percentage adsorption with the divalent exchanged SMT has also been observed. Cu^{2+} exchanged SMT enhance the adsorbing ability of the both SMT (i.e. ~39 and ~26 % for MMT-Cu^{2+} and BNT-Cu^{2+} respectively) while in both the cases for Mg^{2+} exchanged SMT has apparently decreased (~24 and ~34 % respectively) the affinity of Leu for SMT(Table 1). SEM images of BNT-Cu^{2+} and its adduct with Leu are shown Fig. 2, clearly indicating the surface morphology of the BNT-Cu^{2+}, as Leu has finely covered its surface after adsorption (Fig. 2b).

Vibrational studies using FTIR were used to analyse the possible sites of interaction. FTIR spectra of Leu, MMT and Leu–MMT complex are provided in Fig. 3. The characteristic infrared spectral frequencies have been summarized in supplementary material (S.t.1).The bands at ~1035 $cm^{-1}$and ~3622 cm^{-1}corresponds to Si–O–Si stretching vibration and –OH stretching, respectively, while 3431 cm^{-1} and 1638 cm^{-1} are from the stretching and bending vibration of adsorbed water in MMT (Fig. 3b). In the case of Leu, no band corresponds to $-NH_2$ bending and absence of protonated –COOH (i.e. –OH stretching at ~3500 cm^{-1}, –C=O stretching at ~1700−1780 cm-1, –OH bending at ~1200–1300 cm-1), confirming amino group is protonated ($—NH_3^+$) while carboxylic group is deprotonated ($—COO^-$) (i.e. zwitter-ion) (Fig. 3a). Polymerization of Leu is not found in starting adsorbate as well as adsorption adduct as no peak correspond to amide (1640 cm^{-1} and 1550 cm-1) is found (Fig. 3a, c).

A band appears at 3073.17 cm^{-1} which is assigned to an leu-NH_3^+stretch (Fig. 3a) which changes to 3121.95 cm^{-1} after adsorption. Further, shift in the frequencies (i.e.$—COO^-$ symmetric stretching, $—COO^-$ anti-symmetric stretching) also suggest possible interaction via the COO^- moiety of Leu. A lower shift in the vibrational frequencies in adsorption adduct (Leu+MMT) also indicates, weak forces are applied in the course of surface interaction (Stevens and Anderson 1996).

Table 1 Percentage binding and Langmuir isotherm parameters of Leu on the surface of different adsorbents

Adsorbent	Binding (%)	Langmuir Parameters		
	Leu	X_m	K_L	r^2
BNT	24.72	28.83	62414.97	0.99
BNT-Mg^{2+}	24.62	28.75	62417.36	0.99
BNT-Ca^{2+}	25.98	30.36	62421.32	0.99
BNT-Cu^{2+}	26.39	32.32	62417.19	0.99
MMT	34.75	42.24	62422.88	0.99
MMT-Mg^{2+}	34.15	41.84	62413.69	1.00
MMT-Ca^{2+}	37.73	45.69	62421.28	0.99
MMT-Cu^{2+}	39.61	48.48	62410.89	0.99

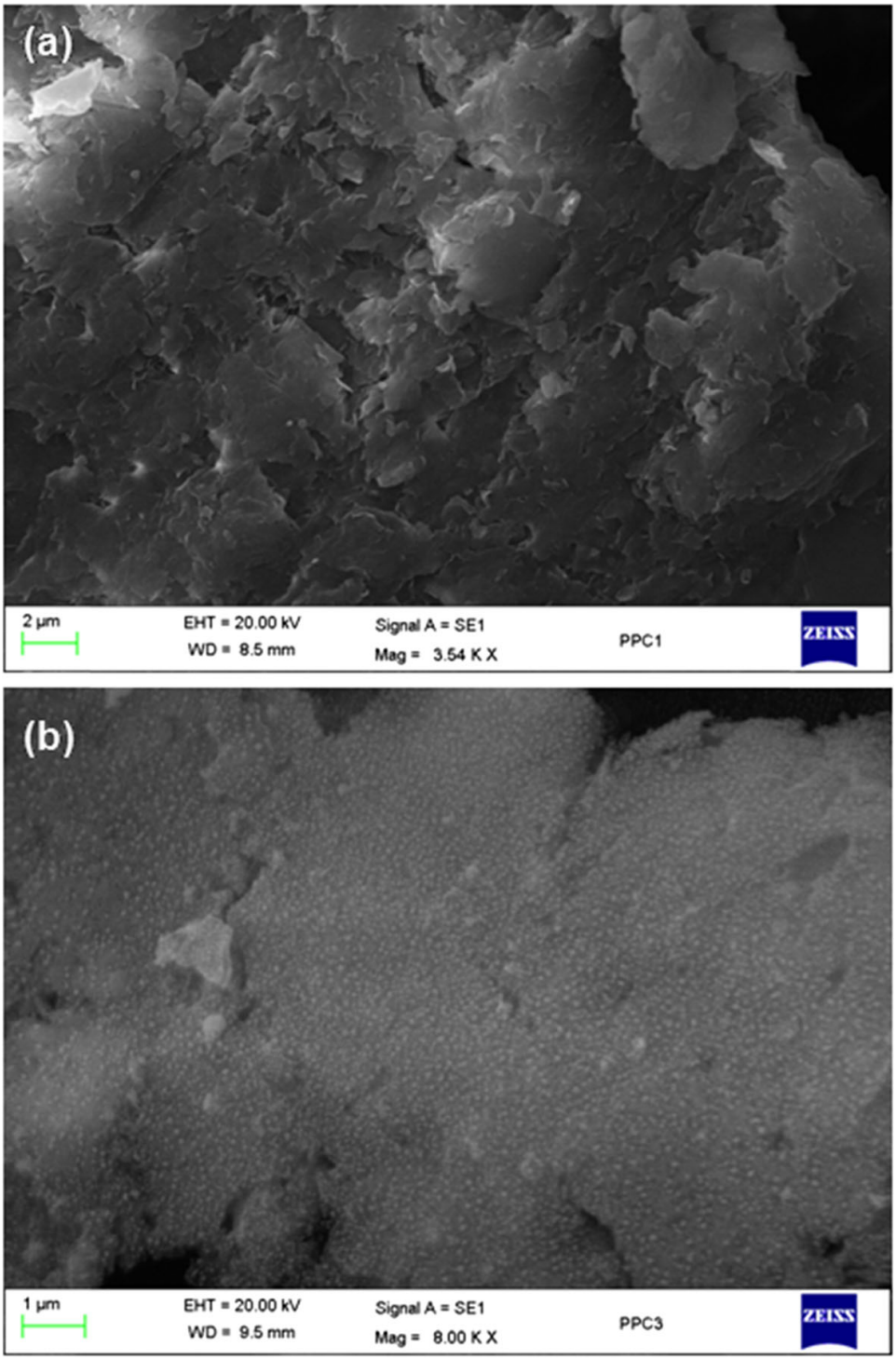


Fig. 2 SEM images of BNT-Cu^{2+} (**a**), and (**b**) adsorption adduct of Leu with BNT-Cu^{2+}

A proton sandwiched between two electronegative atoms (i.e. N of —$NH_3^+$and O of ─Si─O^- and ─Al─O^- of SMT) after adsorption require more energy to vibrate and thus, a shift in the hydrogen bond donor (—NH_3^+) as well as acceptor (—COO^-) sites of the zwitterion of AAs after adsorption on mineral surface found to be the crucial forces to stabilize AAs on SMT (Fig. 3 and S.t.1). The protonated amino group is an excellent hydrogen bond donor (Kollar et al. 2003; Fuchida et al. 2014; Marsh and Donohue 1967). The XRD of adsorption adducts have also, shown a shift in the basal spacing (001). The basal spacing of MMT is 15.56 Å which increases upto 17.18 Å after adsorption (Supplementary material S.4-

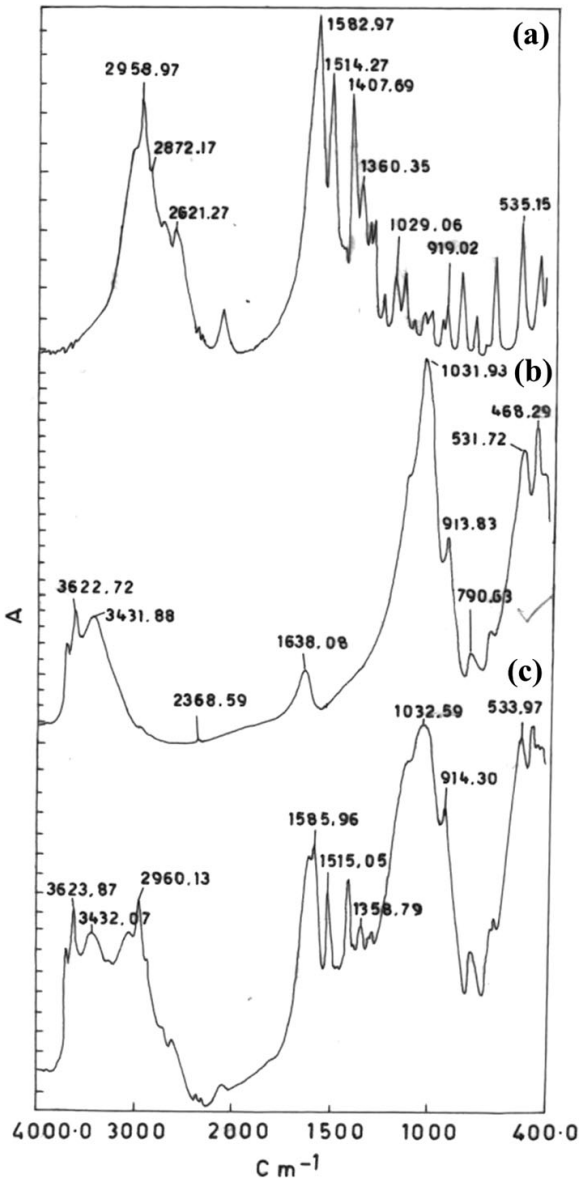


Fig. 3 IR spectra of (**a**) Leu, (**b**) MMT, and (**c**) adsorption adduct, Leu+MMT after adsorption at pH 6.0±0.1

S.6). The intercalation of AAs between the layers of SMT during the adsorption is more prominent than interaction on surfaces and edges of the SMT (Friebele et al. 1980) and an increment in the basal spacing (001) indicates towards the interlayer adsorption where AAs occupy the interlayer space of SMTs because AAs adsorbed on the surface may easily desorbs when SMTs are washed with $_{td}H_20$, prior to the FTIR and XRD analysis, due to the low activation energy of desorption (Lambert 2008). This interlayer acquisition of AAs in SMT may occur via cation exchange (Greenland et al. 1965) which has recently been shown by Fuchida et al. 2014.

Considering the SMT surface, previous reports have also shown surface interactions are largely dependant on interlayer cations, hydrogen bonding, physical forces and the clay's surface charge. (Ponnamperuma et al. 1982; Greenland et al. 1965) while net electric charge of the AAs, which depend on the pH in both the bulk solution and in the proximity of the clay layer or interlayer spaces, is another factor deciding adsorption (Prabhakar et al. 2007). Further, no decomposition or oligomerization of Leu was observed in our experiments.

Computational study

Complexation modes between M^{2+} and Leu with attractive interactions between electron rich (─COO^-) and electron deficient sites (M^{2+}) have been considered for zwitterionic amino acids as reported by Remko et al. (2011) and Marino et al.(2001).The respective atoms of zwitterionic Leu are numbered (not IUPAC) according to the scheme presented in Fig. 4a, and optimized geometric parameters of the Leu considered for theoretical studies along with hydrated or non-hydrated metal complexes Leu—$M^{2+}\times(H_2O)_n$,(where n=0, 2 or 4) are provided in supplementary material S.t.2-S.t.4. All of the output files can also be made available by contacting the corresponding author directly. For the hydrated system Leu—M^{2+}, metal ions are coordinated with two water molecules (Figs. 5a, c and e) using a supermolecule modelling of hydration while other two water molecules were placed near the polar part of the AA (Figs. 5 b, d and f), where the water molecules form intermolecular hydrogen bonds with the proton donor -NH_3^+ group (Remko et al. 2011).

The bond lengths, bond angles and dihedral angles of the computed Leu—M^{2+}complexesaregiven in Table 2, while calculated net charges on the metal, oxygen atoms, nitrogen atoms, dipole moment, total energies, relative energies including ZPE and thermodynamic parameters of nine different complexes of Leu—$M^{2+}\times(H_2O)_n$($M=Cu^{2+}$, Mg^{2+}, Ca^{2+} and n=0, 2 and 4) are summarized in Tables 3 and 4.

Preliminary computations of Leu suggest it forms stable complexes with divalent metal cations that may form in two different ways. First, binding of M^{2+} with the O3 of carboxylic acid group and the N5 of amine in neutral AA while, another, its interaction with both the oxygen (O3 and O4 ends) of the carboxylate ion of the zwitterionic form of the AA. The bicoordinated complexes of M^{2+}via simultaneous interactions of carbonyl oxygen and amine nitrogen is not possible as the lone pair of electron have already been associated with the proton (H^+) liberated from –COOH group in forming the zwitterion. The zwitterionic form of the AA may be stabilized by the presence of counter-ions and/or water molecules (Remko et al. 2011; Bush et al 2008). Further, harmonic vibrational frequencies of the optimized species have shown that all nine complexes are in their minimum energy state as no imaginary frequencies have been observed in their respective frequency calculations using the level of quantum theory as described in experimental section.

The selected geometrical changes for the Leu—M^{2+} complexes indicate that the O(4) ····M(9) bonds are always shorter (0.025 Å, 0.041 Å and 0.032 Å for Mg^{2+}, Ca^{2+} and Cu^{2+}

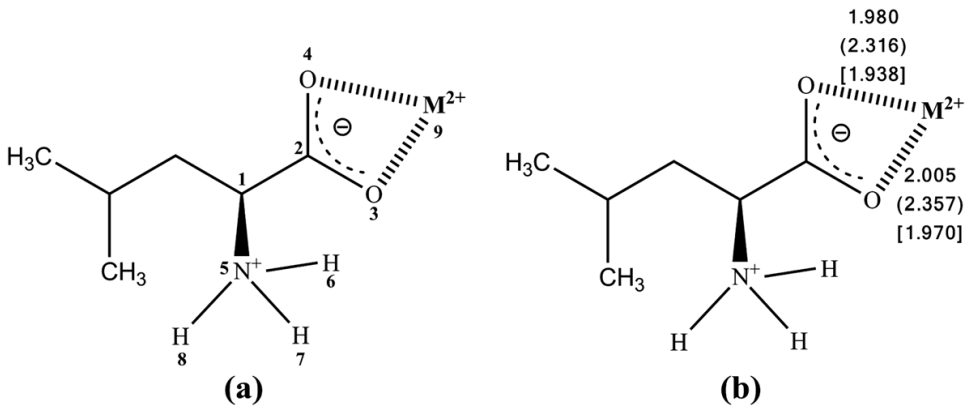


Fig. 4 Chemical structures of (**a**) Leu with numbering, and (**b**) B3LYP/631+G(d,p) optimized distances (Å) for different bicoordinated Leu—M^{2+} complexes with Mg^{2+}, Ca^{2+} and Cu^{2+} reported without enclosures, within parentheses, and with square brackets, respectively

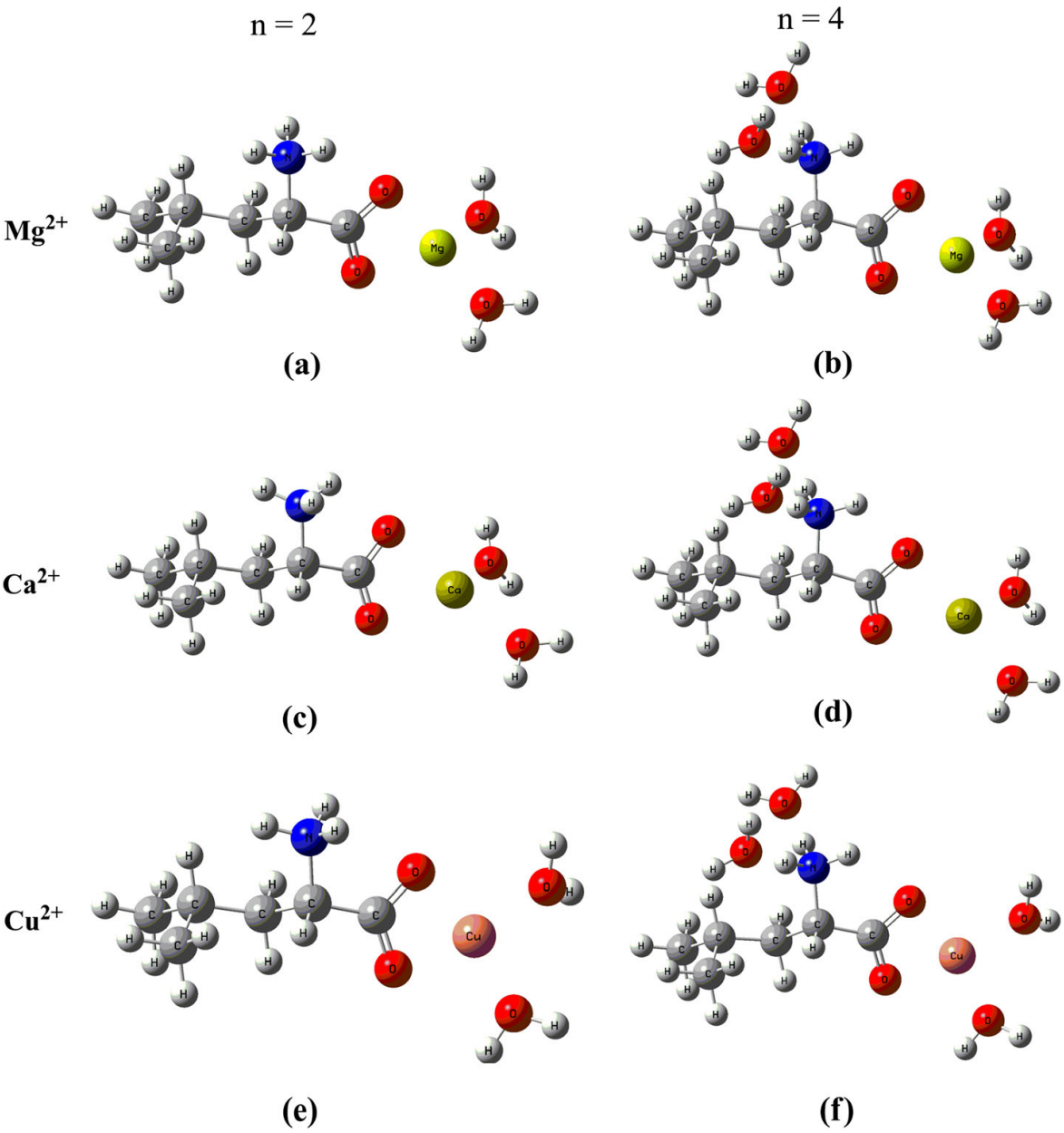


Fig. 5 Overall structure of the B3LYP/631+G(d,p) optimized complexes of Leu with Mg^{2+} (**a** and **b**), Ca^{2+} (**c** and **d**) and Cu^{2+} (**e** and **f**) in presence of water molecules (n=2 or 4)

respectively) than O(3)···M(9) bonds, indicating that these metal ions form stronger bonds with O(4) which are farther away from the ammonium group ($-NH_3^+$) of Leu. The equilibrium distance of O(4)···M(9) and O(3)···M(9) increases in the order: $Cu^{2+}<Mg^{2+}<Ca^{2+}$ (Fig. 4b). The M^{2+} in both the non-hydrated (Leu—M^{2+}) as well as hydrated (Leu—$M^{2+}\times(H_2O)_n$, n=2 or 4) lie in the same plane of carboxylate groups of Leu (Fig. 5).

Table 4 summarizes the different energies and thermodynamic parameters which reflect that the stability of Leu-M^{2+}can be substantially modified by the specific metal ion and hydration. Energy parameters were calculated using eq. 2 and all nine complexes studied were found to form exothermically (Table 4). Leu—$Cu^{2+}$was found to be the most stable complex while its hydration further increases its stability with a maximum enthalpy of -386.28 kcal/mol for Leu—$Cu^{2+}\times(H_2O)_4$, among all nine complexes studied(Table 4). Metal ion affinities as a negative enthalpy variation (ΔH) for Leu—M^{2+}complexes has found minimum with Ca^{2+}(Table 4). Gibbs energies, which measure the tendency to associate in real molecular

Table 2 B3LYP/6-31+G(d,p) optimized relevant bond lengths (Å), bond angles (deg) and dihedral angles (deg) for Leu—M^{2+} complex

Parameters	Leu−M^{2+} complex		
	Mg^{2+}	Ca^{2+}	Cu^{2+}
d[C(1)−C(2)]	1.51555	1.52482	1.50402
d[C(2)−O(3)]	1.28631	1.28202	1.28089
d[C(2)−O(4)]	1.27438	1.26606	1.28120
d[C(1)−N(5)]	1.52211	1.52463	1.52208
d[N(5)−H(6)]	1.02804	1.02950	1.02931
d[N(5)−H(7)]	1.02694	1.02558	1.02866
d[N(5)−H(8)]	1.02918	1.02819	1.03031
d[O(3)−H(6)]	2.51355	2.21961	2.69964
d[O(3)····M(9)]	2.00456	2.35692	1.97009
d[O(4) ····M(9)]	1.97993	2.31611	1.93847
Θ[C(1) –C(2)−O(3)]	121.11858	119.50021	124.50286
Θ[C(2) –C(1)−N(5)]	108.92713	108.00197	109.60729
Θ[O(3) –C(2)−O(4)]	119.02753	121.72136	113.79321
Θ[C(1) –O(4) ····M(9)]	120.48338	125.56861	122.20556
Θ[C(1) –O(3) ····M(9)]	118.24496	122.69798	119.25686
Θ[C(2) –O(3) ····M(9)]	86.16892	89.57749	89.09900
Θ[O(3)····M(9)····O(4)]	67.25545	56.87553	66.60602
ϕ[O(3)−C(2)−C(1)−N(5)]	−8.89928	−12.30457	−8.77769
ϕ[C(1)−C(2)−O(4)···· M(9)]	178.73075	178.25947	179.10865
ϕ[C(1)−C(2)−O(3)···· M(9)]	−178.71686	−178.25057	−179.06574

complexes, are found to be negative and have energies from -114.27 to -342.61 kcal/mol (Table 4). The energetic differentiation of all the complexes studied based on the type of different metal ions (Mg^{2+}, Ca^{2+} and Cu^{2+}) with Leu, a decreasing order of Gibbs energy is Leu—Cu^{2+}>Leu—Mg^{2+}>Leu—Ca^{2+}has been observed. The ΔG for the non-hydrated leu-M^{2+}complexes were calculated to be -173.09 k.cal/mol, -114.27 k.cal/mol and -251.34 k.cal/

Table 3 B3LYP/6-31+G(dp) calculated net charge of metal (q_{M+}), two oxygen ($q_{O(3)}$ and $q_{O(4)}$), nitrogen ($q_{N(5)}$) in |e|, dipole moments (μ in Debye) and amount of metal ion charge transfer (CT) for different Leu—M^{2+}×$(H_2O)_n$, (n=0, 2 and 4) systems calculated at 298.15 K and 1 atm. pressure

System	Species		q_{M+}	q_{Ons}	$q_{O(4)}$	$q_{N(5)}$	μ	CT/e
(Leu····M^{2+})×$(H_2O)_n$	M=Mg^{2+}	n=0	1.822	−0.834	−0.789	−0.818	15.25	0.178
		n=2	1.790	−0.834	−0.784	−0.818	12.92	0.210
		n=4	1.787	−0.840	−0.802	−0.855	13.11	0.213
	M=Ca^{2+}	n=0	1.924	−0.859	−0.808	−0.821	16.82	0.076
		n=2	1.907	−0.850	−0.793	−0.825	13.68	0.093
		n=4	1.902	−0.854	−0.813	−0.855	14.39	0.098
	M=Cu^{2+}	n=0	1.332	−0.624	−0.576	−0.821	6.60	0.668
		n=2	1.401	−0.714	−0.667	−0.818	7.17	0.599
		n=4	1.384	−0.712	−0.675	−0.854	7.58	0.616

Table 4 B3LYP/631+G(d,p) calculated absolute energy (E in au), energy change (ΔE in k.cal/mol), metal ion affinities (ΔH in k.cal/mol), Gibbs free energies (ΔG in k.cal/mol) and entropies (ΔS in cal/mol.K) for different Leu—$M^{2+}\times(H_2O)_n$, (n=0, 2 and 4) systems calculated at 298.15 K and 1 atm pressure

Parameters	$(Leu^{\cdots}M^{2+})\times(H_2O)_n$								
	$M=Mg^{2+}$			$M=Ca^{2+}$			$M=Cu^{2+}$		
	n=0	*n*=2	*n*=4	*n*=0	*n*=2	*n*=4	*n*=0	*n*=2	*n*=4
-E	641.25	794.28	947.20	1118.79	1271.78	1424.70	2081.39	2234.42	2387.34
-ΔE	182.73	284.14	321.72	123.74	199.61	234.69	260.10	356.99	394.52
-ΔH	181.11	278.72	313.24	122.27	194.60	226.61	259.23	349.62	386.28
-ΔG	173.09	252.94	271.25	114.27	170.81	187.68	251.34	323.48	342.61
-ΔS	26.89	86.46	140.86	26.83	79.81	130.57	26.47	87.66	146.45

mol for Mg^{2+}, Ca^{2+}, and Cu^{2+} complexes respectively. To analyze the affinity of hydration with the Leu—M^{2+} complexes, first two water molecules were coordinated to M^{2+} ions (Figs. 5a, c and e), which lowered the ΔG values for the complexes with Mg^{2+}, Ca^{2+}, and Cu^{2+} by-79.85 k.cal/mol, -56.54 k.cal/mol and -72.14 k.cal/mol respectively (Table 4). Addition of another two water molecules to give an Leu—$M^{2+}\times(H_2O)_4$ complex further lowered the ΔG values by -18.31 k.cal/mol (Fig. 5b), -16.87 k.cal/mol (Fig. 5d) and -19.13 k.cal/mol (Fig. 5f) for Mg^{2+}, Ca^{2+}, and Cu^{2+} complexes respectively.

The net atomic charges on the different atoms of the complexes studied, have been investigated by the natural population analysis using the NBO program (Reed and Weinhold 1985; Reed et al. 1985; Glendening et al. NBO program, Version 3.1). The amounts of charge transfer (CT) for the M^{2+} along with dipole moments are summarized in Table 3. The stronger bond formation between $O4^{\cdots}M^{2+}$of different Leu—$M^{2+}\times(H_2O)_n$ complexes can further be analyzed through the charges on O4 along with $O4^{\cdots}M^{2+}$distance (Å) in comparison to charge and bond distance (Å) of O3 and $O3^{\cdots}M^{2+}$ respectively. The charge on O4 of different complexes is always in lower side than O3 which suggest the transfer of more of it charge towards M^{2+} from this end contrary to O3. The magnitude of CT contribution in Leu—$Ca^{2+}$is much lower (0.076 to 0.098*e*) for its non-hydrated as well as hydrated complexes while CT for all Leu—Cu^{2+} complexes are maximum from 0.559 to 0.668*e* (Table 3). The magnitudes of charge transfer on the metal ions studied are in good agreement with their metal ion affinities (Table 3 and 4). The interactions of glycine with 15 different metals at theoretical level, also showed the magnitude of CT correlates well with metal ion interaction energies (Hoyau et al. 2001).

Looking at the bigger picture, the experimentally studied adsorption processes of Leu on SMT and their Cu^{2+} and Ca^{2+} ion exchanged counterparts, essentially represents the enhanced affinity of biomolecules on metal-exchanged SMT, which have been suggested to have played a role during chemical evolution and the origin of life (Lambert 2008). The present study shows significant affinity of non-polar amino acid Leu with M^{2+} ion exchanged SMT surfaces, specially Cu^{2+} exchanged SMT, which might be helpful to give insight into the higher percentage of neutral and non-polar AAs in proteins of present day living systems as modern proteins comprise of total 74 % non polar amino acids (Zaia 2004). Metal exchanged minerals could possibly help to accumulate the non-polar amino acids via adsorption on their surfaces which may, facilitate their polymerization to get complex proteins. Further, this study may help

to understand the specific sites of attachments of amino acids, peptides or proteins onto different field of adsorption (peptide sequencing identification, nano technology, drug targeting/delivery and chemical evolution) (Roman et al. 2006) and the physico-chemical parameters (i.e. change in bond length, bond angles, electron density, charge and thermodynamic parameters) of M^{2+}—AAs complexes may also help in better understanding of the biochemistry of living systems and also serve as a simpler model for the AAs—M^{2+} interaction found in the more complex protein, enzyme-metal systems related to living systems such as Cu^{+} coordinated with peptides (viz. glycylglycine) has been reported to stabilize this peptide structure (Constantino et al. 2005).

Conclusion

The interaction of the Leu on BNT/MMT and its divalent metal ion (Mg^{2+}, Ca^{2+} and Cu^{2+}) exchanged forms was investigated with respect to the optimal conditions of pH, time and concentration at 298 K. Adsorption parameters such as percentage binding show that Leu is adsorbed more on the MMT than BNT. Further, Cu^{2+} exchanged BNT/MMT adsorbs Leu the most strongly, while Mg^{2+}exchanged SMTs showed the weakest sorption. Experimentally, interaction of biomolecules with metal exchanged clays and theoretically determined physico-chemical characteristics of different metal ion-biomolecular interactions, may shed light on the prebiotic formation of macromolecules and their significance in living systems such as affinity of Cu^{2+} for the biomolecular selectivity for the enhancement of peptide formation (Sakata et al. 2014; Constantino et al. 2005). A better surface interaction of AA in presence of Cu^{2+} incorporated minerals might possibly enhance the proximity effect of functional groups to form peptides. Moreover, it may be concluded that in prebiotic era of the earth adsorption processes involving metal ions might have played a key role for local concentration of biomolecules and thus facilitated further condensation and polymerization on their surfaces into more complex molecules essential for life.

Acknowledgment Authors are grateful to two anonymous reviewers for the major suggestion and corrections. The authors are also thankful to Head of the Chemistry Department and Prof. Sanjay Pant Photophysics Laboratory, Head of Physics Department, DSB Campus, Kumaun University, Nainital 263002, India for providing necessary research facilities. Prof. Sanjay Pant is specially acknowledged for providing computational facility of Gaussian 03w, along with other research facilities. Authors (PA and NP) acknowledge the UGC-RFMS (Sanction No:F, 5-53/2007 BSR), New Delhi, India for financial assistance.

References

Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648

Bernal JD (1951) The Physical basis of life.ProcPhysSocLond A. 62:537-558

Bush MF, Oomens J, Saykally RJ, Williams ER (2008) Effects of alkaline earth metal ion complexation on amino acid zwitterion stability: results from infrared action spectroscopy. J Am Chem Soc 130:6463–6471

Caglar B, Afsin B, Tabak A, Eren E (2009) Characterization of the cation-exchanged bentonites by XRPD, ATR, DTA/TG analyses and BET measurement. Chem Eng J 149:242–248

Carr MH, Head (III) JW (2010) Geological history of mars. Earth Planet Sci Lett 294:185–203

Cleaves HJ(II), Scott AM, Hill FC, Leszczynski J, Sahai N, Hazen R (2012) Mineral–organic interfacial processes: potential roles in the origins of life. Chem Soc Rev 41:5502–5525

Constantino E, Rimola A, Santiago LR, Sodupe M (2005) Coordination properties of glycylglycine to Cu^{+}, Ni^{+} and Co^{+}: influence of metal cation electronic configuration. New J Chem 29:1585–1593

Ferris JP, Hill AR Jr, Liu R, Orgel LE (1996) Synthesis of long prebiotic oligomers on mineral surfaces. Nature 381:59–61

Friebele E, Shimoyama A, Ponnamperuma C (1980) Adsorption of protein and non-protein amino acids on a clay minerals: a possible role of selection in chemical evolution. J MolEvol 16:269–278

Frisch MJ, Trucks GW, Schlegel HB, Scuseria, Robb MA, Cheeseman JR, Montgomery JA, VrevenJrT, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA,Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C,Ochterski JK, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA(2003) Gaussian 03, Revision A.1, Gaussian, Inc., Pittsburgh, PA

Fuchida S, Masuda H, Shinoda K (2014) Peptide formation mechanism on montmorillonite under thermal conditions. Orig Life EvolBiosph 44(1):13–28

Glendening ED, Reed AE, Carpenter JE, Weinhold F. The NBO program, Version 3.1

Greenland DJ, Laby RH, Quirk JP (1965) Adsorption of amino-acids and peptides by montmorillonite and illite. Part 1. Cation exchange and proton transfer. Trans Faraday Soc 61:2013–2023

Gururani K, Pant CK, Pandey N, Pandey P (2012a) Heat induced formation of peptides from reaction mixture of glycine-glutiamic acid and glycine-leucine in presence and absence of montmorillonite clay with or without metal ions under wetting drying cycles of primitive Earth. Int J Scie Techno Res 1(8):159–163

Gururani K, Pant CK, Pandey P, Pandey N (2012b) Heat induced synthesis of amino acids from reaction system comprised of acetylene, ammonia and water vapour in presence and absence of metal oxide (silica and alumina) under wetting-drying condition of primitive Earth. Asian J Res Chem 5(9):1150–1155

Hohenberg P, Kohn W (1964) Inhomogeneous electron Gas. Phys Rev 136:B864–B871

Hoyau S, Pelicier JP, Rogalewiez F, Hoppilliard Y, Ohanessian G (2001) Complexation of glycine by atomic metal cations in the gas phase. Eur J Mass Spectrom 7:303–311

Kalra S, Pant CK, Pathak HD, Mehata MS (2003) Studies on the adsorption of peptides of glycine/alanine on montmorillonite clay with or without co-ordinated divalent cations. Colloids and Surfaces A: Physicochem Eng Aspects 212:43–50

Kebarle P, Hogg AM (1965) Mass-spectrometric study of ions at near atmospheric pressures. I Ionic Polym Ethyl J Chem Phys 42(2):668–674

Kollar T, Palinko I, Konya Z, Kiricsi I (2003) Intercalating amino acid guests into montmorillonite host. J MolStruct 651:335–340

Lahav N, Chang S (1976) The possible role of solid surface area in condensation reactions during chemical evolution: reevaluation. J MolEvol 8:357–380

Lambert JF (2008) Adsorption and polymerization of amino acids on mineral surfaces: a review. Orig Life Evol Biosph 38:211–242

Lee C, Yang W, Parr RG (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785

Lucrezia DD, Anella F, Chiarabelli C (2007) Question 5: on the chemical reality of the RNA world. Orig Life EvolBiosph 37:379–385

Marino T, Russo N, Toscano M (2001) Potential energy surfaces for the gas-phase interaction between α-alanine and alkali metal ions (Li+, Na+, K+): a density function study. Inorg Chem 40:6439–6443

Marsh RE, Donohue J (1967) Crystal structure studies of amino acids and peptides. AdvProtein Chem 22:235–256

Marshall-Bowman K, Ohara S, Sverjensky DA, Hazen RM, Cleaves HJ (2010) Catalytic peptide hydrolysis by mineral surface: implications for prebiotic chemistry. Geochim Cosmochim Acta 74:5852–5861

Merz KM, Aguiar EC, da Salvia JBP (2014) Adenine formation without HCN. J Phys Chem A 118(20):3637–3644

Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. Crit Rev Biochem Mol Biol 39:99–123

Pandey P, Pant CK, Gururani K, Arora P, Kumar S, Sharma Y, Pathak HD, Mehata MS (2013) Surface interaction of L-alanine on hematite: anastrobiological implication. Orig Life Evol Biosph 43:331–339

Pant CK, Lata H, Pathak HD, Mehata MS (2009) Heat-initiated prebiotic formation of peptides from glycine/aspartic acid and glycine/valine in aqueous environment and clay suspension. Int J Astrobiol 8(2):107–115

Pizzarello S (2009) The cosmochemical record of carbonaceous meteorites:an evolutionary story. J Mex Chem Soc 53(4):253–260

Ponnamperuma C, Shimoyama A, Friebele E (1982) Clay and origin of life. Orig Life Evol Biosph 12:9–40
Prabhakar A, Cuadros J, Sephton M, Dubbiam D, Coles BJ, Weiss D (2007) Adsorption of L-lysine on montmorillonite. Colloids Surf A Physicochem Eng Asp 307:142–149
Reed AE, Weinhold FA (1985) Natural localized molecular orbitals. J Chem Phys 83:1736–1740
Reed AE, Weinstock RB, Weinhold FA (1985) Natural population analysis. J Chem Phys 83:735–746
Remko M, Rode BM (2000) Bivalent cation binding effect on formation of the peptide bond. Chem Phys Lett 316:489–494
Remko M, Fitz D, Broer R, Rode BM (2011) Effect of metal ions (Ni^{2+}, Cu^{2+} and Zn^{2+}) and water coordination on the structure of L-phenylalanine, L-tyrosine, L-tryptophan and their zwitterionic forms. J Mol Model 17: 3117–3128
Rode BM, Bujdak J, Eder AH (1993) The role of inorganic substances in the chemical evolution of peptides on Earth. Trends Inorg Chem 3:45–62
Rodgers MT, Armentrout PB (2000) Noncovalent interactions of nucleic acid bases (uracil, thymine, and adenine) with alkali metal ions. Threshold collision-induced dissociation and theoretical studies. J Am Chem Soc 122(35):8548–8558
Roman T, Dino WA, Nakanishi H, Kasai H (2006) Glycine adsorption of Single-walled carbon nanotubes. Thin Solid Films 509:218–222
Sakata K, Yabuta H, Kondo T (2014) Effects of metal ions and p H on the formation and decomposition rates of di- and tri-peptides in aqueous solution. Geochem J 48(2):219–230
Shapiro R (1999) Prebiotic cytosine synthesis: a critical analysis and implications for the origin of life. Proc Natl Acad Sci U S A 96:4396–4401
Stephenson JD, Hallis LJ, Nagashima K, Freeland S (2013) Boron enrichment in martian clay. PLoS ONE 8(6): 1–4
Stevens JJ, Anderson SJ (1996) An FRIR study of water sorption on TMA- and TMPA-montmorillonites. Clay Clay Miner 44(1):142–150
Stribling R, Miller SL (1987) Energy yields for hydrogen cyanide and formaldehyde syntheses: the HCN and amino acid concentrations in the primitive ocean. Orig Life Evol Biosph 17:261–273
Zaia DAM (2004) A review of adsorption of amino acids on minerals: was it important for origin of life? Amino Acids 27:113–118
Zaia DAM (2012) Adsorption of amino acids and nucleic acid bases onto minerals: a few suggestions for prebiotic chemistry experiments. Int J of Astrobiol 11:1–6

Microsyst Technol
DOI 10.1007/s00542-015-2554-z

TECHNICAL PAPER

Analysis of novel transparent gate recessed channel (TGRC) MOSFET for improved analog behaviour

Ajay Kumar[1] · Neha Gupta[1] · Rishu Chaujar[1]

Received: 30 October 2014 / Accepted: 16 April 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In this paper, a novel device structure called transparent gate recessed channel MOSFET (TGRC-MOSFET) is proposed to alleviate the hot carrier effects for the advanced nanometer process. TGRC-MOSFET involving a recessed channel and incorporates indium tin oxide as a transparent gate. TCAD analysis shows that the performance of TGRC-MOSFET surpasses conventional recessed channel (CRC)-MOSFET in terms of high I_{ON}/I_{OFF} ratio and better carrier transport efficiency in comparison to CRC-MOSFET. This simulation divulges the reduction in hot-carrier-effects metrics like electron velocity, electron temperature, potential, and electron mobility. Furthermore, the effect of gate length is observed on the analog behavior of TGRC-MOSFET. All the simulations have been done using DEVEDIT-3D and ATLAS device simulator. The work proposes the novel design for reduced hot carrier and low power switching applications.

1 Introduction

Hasten in semiconductor manufacturing techniques and more and more demand for high speed and more complicated integrated circuits (ICs) have driven the associated metal oxide semiconductor field effect transistor (MOSFET) sizes close to their physical limits. On the other hand, it is difficult to scale-down the supply voltage used to perform these ICs consistently due to compatibility problem with earlier generation circuits, power, noise margin and delay requirements, and not scaling of threshold voltage and subthreshold slope (Torabi et al. 2013; Gautam et al. 2013). While the successive increase in internal electric fields in aggressively scaled MOSFETs comes with the additional ameliorate of increased carrier velocities, and hence increased switching speed, it also presents higher reliability complications for the long period of operation of these devices.

As devices are scaled down, the benefits of higher electric fields saturate while the associated reliability problems get worse. The presence of large electric fields in MOSFETs implies the presence of high energy carriers, referred as "hot-carriers", in such devices (Arora 2007). The presence of this kind of mobile carriers in the oxide creates some physical damages which may change device characteristics for a long time. The aggregation of damage can ultimately be the reason of circuit failure. The aggregation damage of the hot carrier results in deterioration of device behavior (Orouji and Kumar 2005; Pan et al. 1994; Krieger et al. 2001). The need of superior performance of ICs has led to the scaling of MOSFETs down to 30 nm and below.

The hot-carrier dilapidation becomes a stringent limitation to the reliability of 100-nm devices and VLSI packing density (Chaujar et al. 2008a, b). Under the influence of high lateral fields in short-channel MOSFETs, carriers in the channel and pinch-off regions of the transistor reach non-equilibrium energy distributions. The generation of these hot-carriers is the primary source of several reliability problems. Hot-carriers can acquire sufficient energy to

✉ Rishu Chaujar
rishu.phy@dce.edu

Ajay Kumar
ajaydtu@gmail.com

Neha Gupta
nehagupta_dtu@yahoo.com

[1] Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

Published online: 06 May 2015

surmount the energy barrier at the Si–SiO_2 interface or tunnel into the oxide.

From the last several decades, the hot carrier reliability performance in a MOS device has been studied since MOS device is scaled and suffered from hot carrier effect which also pervert analog circuit desideratum. When we scaled the device, the ameliorates of higher electric fields impregnate while associated reliability problems get worse. When electric field will be large in MOSFETs, then carrier will have high energy and these carriers are hot carriers. To abate these hot-carrier effects in CRC-MOSFET, instead of metallic gate we are using indium tin oxide (ITO). ITO (tin-doped indium oxide) is a solid solution of indium oxide (In_2O_3) and tin oxide (SnO_2). It is transparent and colorless in thin layers. ITO is one of the most widely used transparent conducting oxides. By using ITO, we improve I_{ON} which causes the decrease in device power consumption (Salehi and Gholizadeh 2011). The In_2O_3 phase itself contributes free electron for electrical conductivity (Pammi et al. 2011). When tin (Sn) is diffused, then some of the oxygen vacancies may be created by SnO_2 which creates free electrons to enhance the concentration of carriers and hence increase conductivity and decrease the resistivity with temperature.

2 Device structure and its parameters

The simulation device structure i.e. TGRC-MOSFET consists of gate which is made by transparent conducting material ITO as shown in Fig. 1. The total gate length is 30 nm and thickness of oxide is 2.0 nm. In this case, substrate doping is p-type with concentration of 1×10^{16} cm^{-3}; source and drain are n-type with uniform doping profiles is 1×10^{19} cm^{-3}.

All simulations have been performed using ATLAS and DEVEDIT 3D device simulator. In gate electrode, we are replacing metal gate by transparent conducting material ITO, but source electrode and drain electrodes are remain metal. Bias voltage $V_{gs} = 0.7$ V and $V_{ds} = 0.5$ V are applied in all the results except $V_{ds} = 0.8$ V given when calculating DIBL. Design parameters of CRC-MOSFET and TGRC-MOSFET are given in Table 1.

3 Computer simulation results of transparent gate recessed channel MOSFET

The present analysis is carried out for a channel length, $L_G = 30$ nm, uniformly doped source/drain, N_D with doping density of 1×10^{19} cm^{-3}, p-type substrate doping, N_A with a doping density of 1×10^{16} cm^{-3}, SiO_2 thickness, $t_{ox} = 2.0$ nm. The gate work function (Φ_{ITO}) is 4.7 eV.

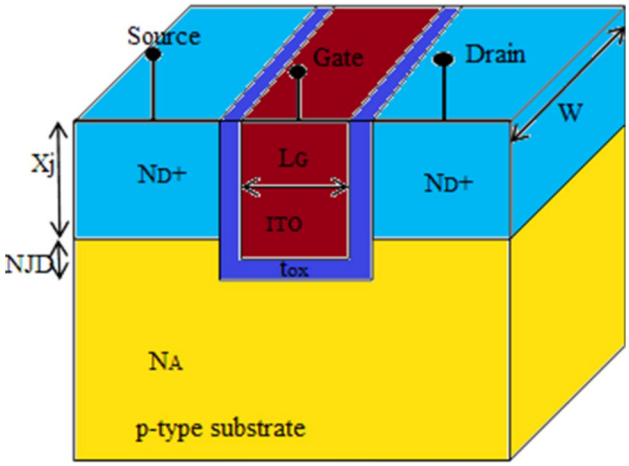


Fig. 1 Simulated device: TGRC-MOSFET

Table 1 Design parameters of CRC-MOSFET and TRGC-MOSFET device designs

Channel length (L_G)	30 nm
Device width	200 nm
Groove depth	38 nm
Source/drain junction depth	30 nm
Negative junction depth (NJD)	10 nm
Substrate doping (N_A)	1×10^{16} cm^{-3}
Source/drain doping (N_D^+)	1×10^{19} cm^{-3}
Physical oxide thickness (t_{ox})	2 nm
Permittivity of SiO_2	$\varepsilon_{ox} = 3.9$
Gate to source voltage (V_{gs})	0.7 V
Drain to source voltage (V_{ds})	0.5 V
Work function for TGRC-MOSFET (Φ_{ITO})	4.7 eV
Work function for CRC-MOSFET (Φ_M)	4.1 eV

3.1 Surface potential, conduction band energy, valence band energy and DIBL

Figure 2a shows the variation of surface potential along the channel for CRC-MOSFET and TGRC-MOSFET at two different drain voltages $V_{ds} = 0.5$ V and 0.8 V.

Figure 2b and c shows the conduction band and valence band energy profiles along the channel length for CRC-MOSFET and TGRC-MOSFET. The lowering of barrier potential at source side indicates that the DIBL effect is less observed in this device, as it is evident from Fig. 2d (Shee et al. 2014; AI-Mistarihi et al. 2013). The improvement in DIBL effect is more in case TGRC-MOSFET, as it incorporates the non-stationary effects such as velocity overshoots. The improvement in DIBL effect is due to the step in conduction band energy profile in the channel

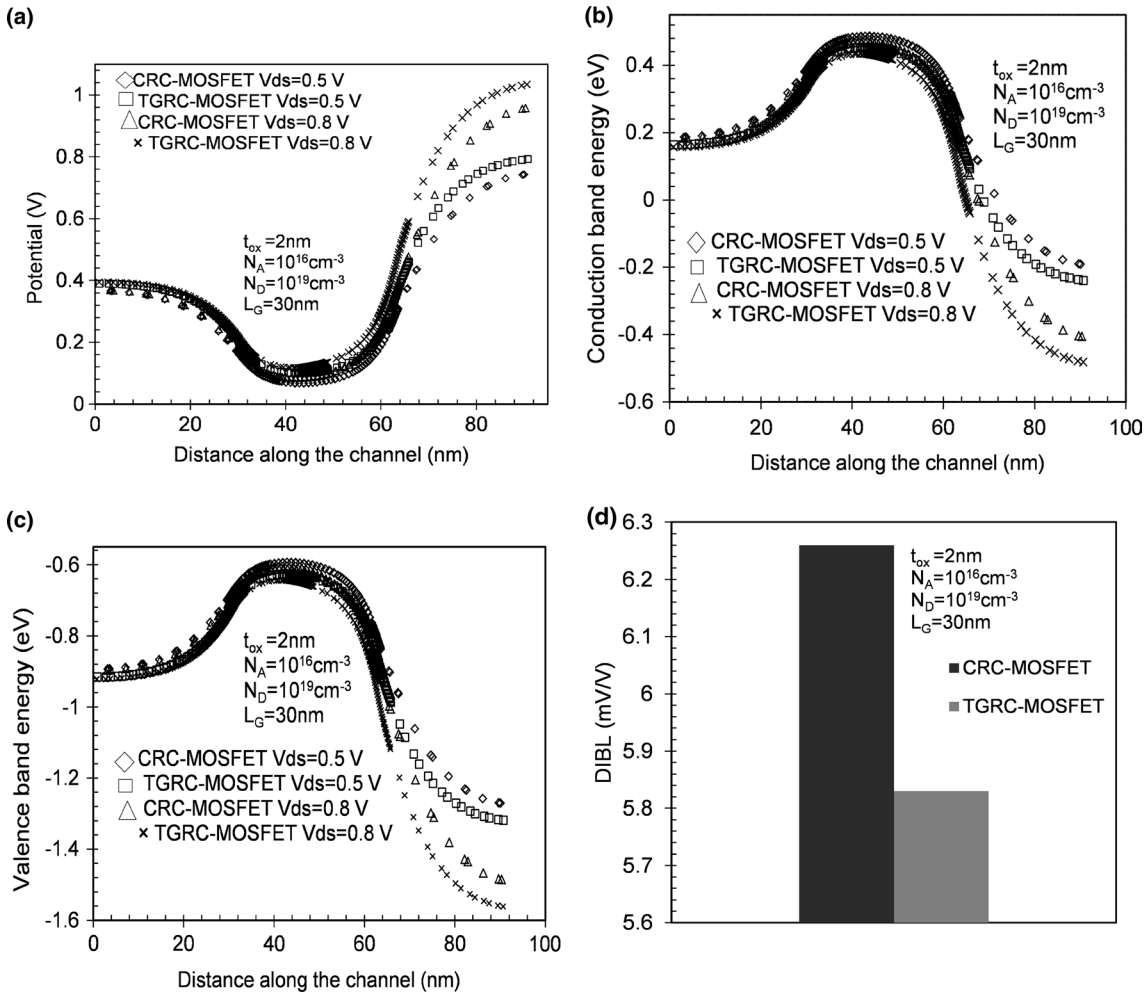


Fig. 2 **a** Surface potential as a function of position along the channel for CRC-MOSFET and TGRC-MOSFET at V_{ds} = 0.5 V and V_{ds} = 0.8 V. **b** Conduction band energy as a function of position along the channel for CRC-MOSFET and TGRC-MOSFET at V_{ds} = 0.5 V and V_{ds} = 0.8 V. **c** Valance band energy as a function of position along the channel for CRC-MOSFET and TGRC-MOSFET at V_{ds} = 0.5 V and V_{ds} = 0.8 V. **d** DIBL for CRC-MOSFET and TGRC-MOSFET at V_{ds} = 0.5 V and V_{ds} = 0.8 V

region, which is resulting from transparent gate. DIBL is more pronounced in conventional MOSFET, which is calculated from the difference of conduction band energy between V_{ds} of 0.8 and 0.5 V. Variation of the bending energy gap due to the threshold voltage variation (ΔV_{th}) is mainly governed by the increase of the drain voltage. Its effect is called the DIBL (AI-Mistarihi et al. 2013; Shee et al. 2014).

3.2 Electric field, electron velocity and ON current

Figure 3 shows the variation of electric field along the channel. It can be seen that in the TGRC-MOSFET, the electric field peak near the source side is higher compared to drain side. The reduction in electric field at the drain end can be interpreted as a reduction in hot carrier effects, lower impact ionization and higher breakdown voltage. Figure 4 shows the variation of electron velocity of TGRC-MOSFET and CRC-MOSFET along the channel. Here the electron velocity is more at source side in TGRC-MOSFET in compare to CRC-MOSFET. As Fig. 4 shows electron velocity at drain side is lower in TGRC-MOSFET in comparison to CRC-MOSFET which reduces hot carrier effects. Figure 5 shows high ON current in TGRC-MOSFET in comparison to CRC-MOSFET which favors its pertinence in switching applications as compared to CRC-MOSFET. In ITO, In_2O_3 contributes many free electrons which increases the electrical conductivity and decreases the resistivity. As a result, the electron velocity near the source side is higher in TGRC-MOSFET which magnifies the source carrier injection into the channel. Thus, the trans-conductance and drive current of TGRC-MOSFET will be more than CRC-MOSFET (Chaujar et al. 2008; Long and Ken 1997).

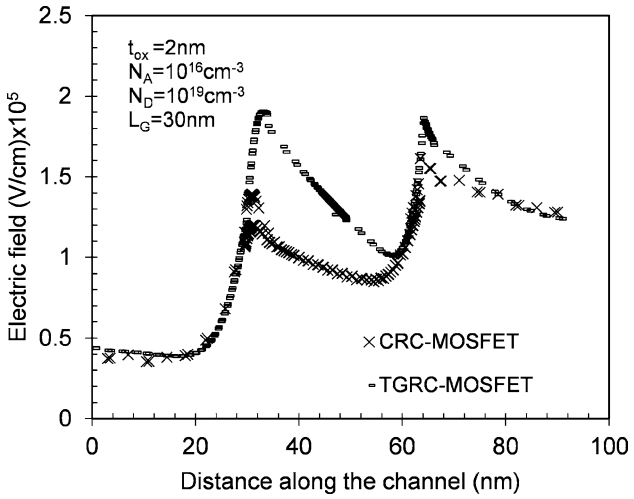


Fig. 3 Electric field as a function of position along the channel for CRC-MOSFET and TGRC-MOSFET at $L_G = 30$ nm, $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V

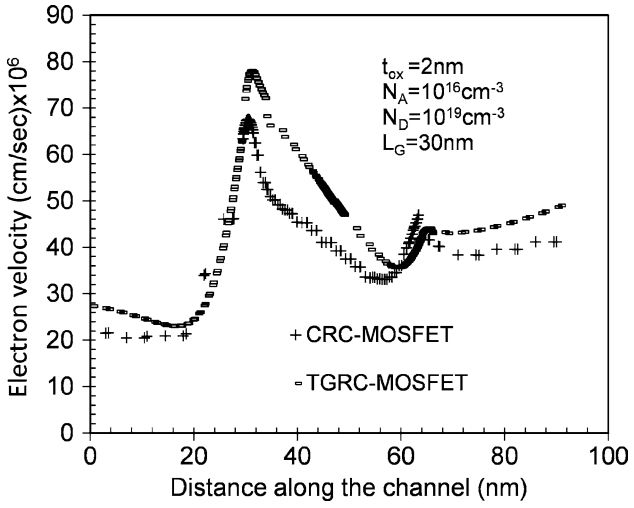


Fig. 4 Electron velocity as a function of position along the channel for CRC-MOSFET and TGRC-MOSFET at $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V

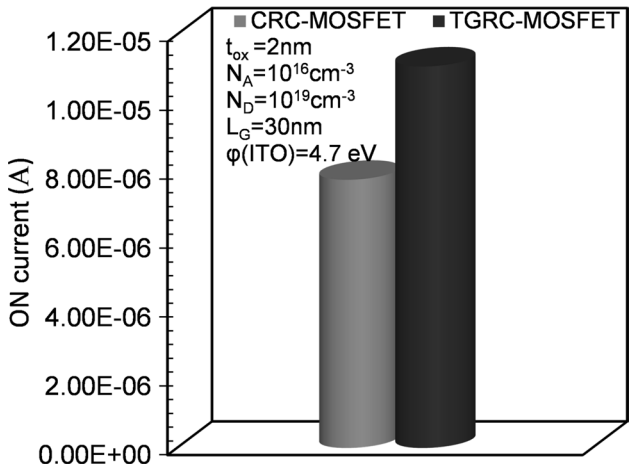


Fig. 5 ON current for CRC-MOSFET and TGRC-MOSFET at $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V

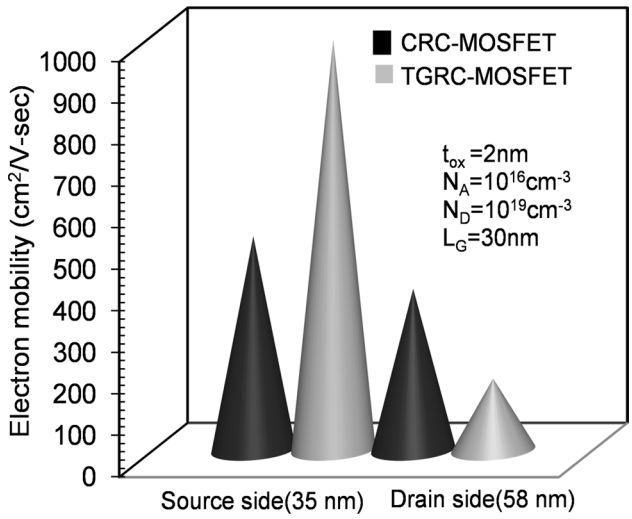


Fig. 6 Electron mobility of TGRC-MOSFET and conventional RC-MOSFET at 35 and 58 nm along the channel at $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V

3.3 Electron mobility and electron temperature

Figure 6 shows the variation of electron mobility of TGRC-MOSFET and CRC-MOSFET along the channel. Here the electron mobility is high at source side in TGRC-MOSFET in comparison to CRC-MOSFET. Figure 6 shows that electron mobility at drain side is higher in CRC-MOSFET in compare to TGRC-MOSFET which reduces hot carrier effects. The mobility of charge carriers in ITO thin films is largely stimulated by four scattering mechanisms: lattice scattering, ionized impurity scattering, neutral impurity scattering, and grain boundary scattering (Thilakan and Kumar 1997). In ITO, tin oxide creates more electrons, thereby increasing the carrier concentration and hence decreasing the mobility at drain side and increasing at source side.

Figure 7 shows the variation of electron temperature across the channel length of MOSFET. In CRC-MOSFET high electron temperature at the drain side ostentation that hot carrier effect is more in CRC-MOSFET as compared to TGRC-MOSFET. When we use ITO as a transparent conducting material, mobility and carrier concentration is directly dependent upon temperature. When carrier concentration increases and mobility decreases, then the increase in electron temperature is less in TGRC-MOSFET

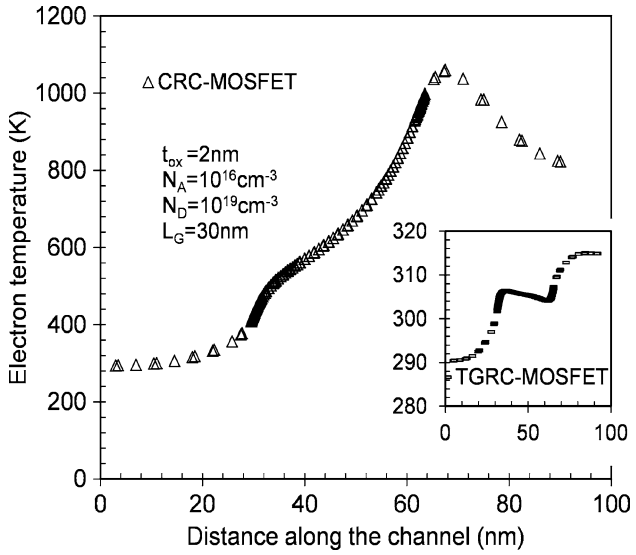


Fig. 7 Electron temperature as a function of position along the channel for CRC-MOSFET at $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V. *Inset* Electron temperature as a function of position along the length for TGRC-MOSFET at $V_{ds} = 0.5$ V and $V_{gs} = 0.7$ V

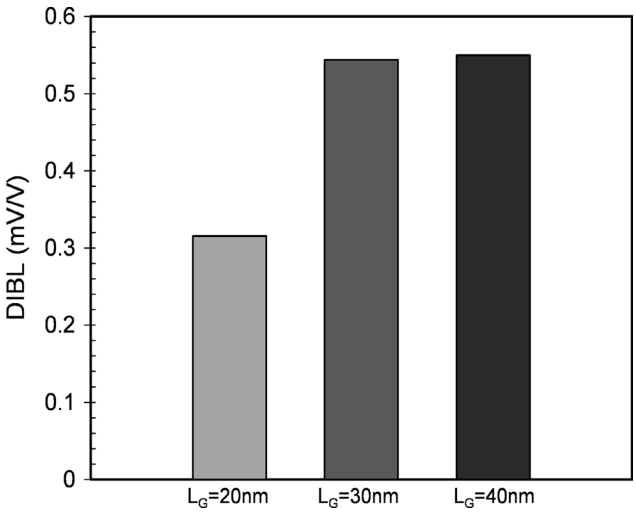


Fig. 8 DIBL variation for TGRC-MOSFET at different gate lengths

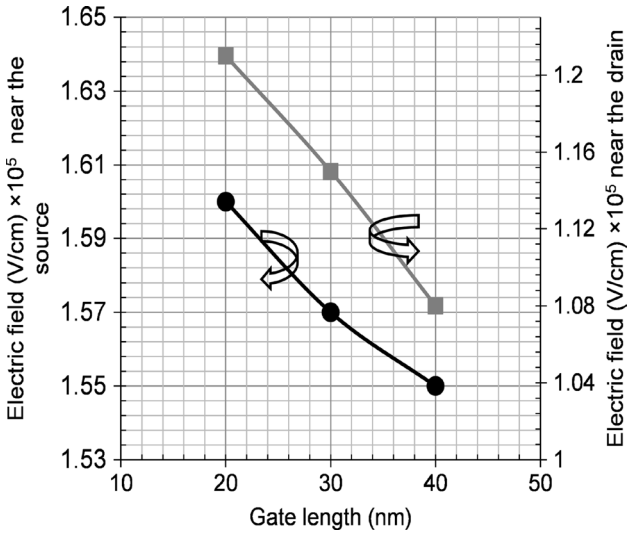


Fig. 9 Electric field variation along the channel for TGRC-MOSFET at different gate lengths

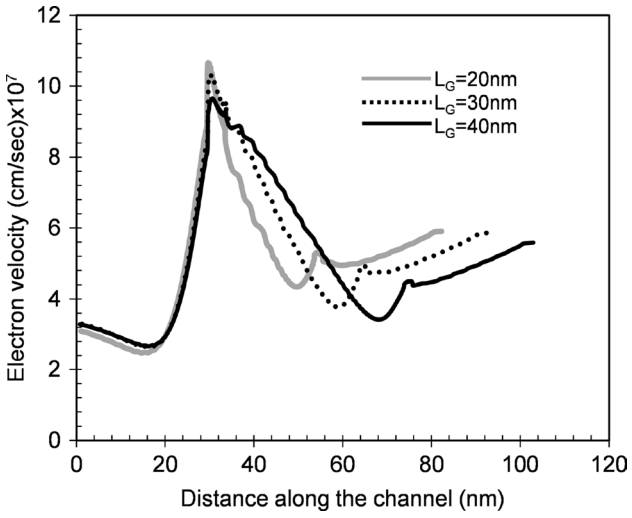


Fig. 10 Electron velocity variation along the channel for TGRC-MOSFET at different gate lengths

compared to CRC-MOSFET. Increase in electron temperature may cause the device damage and hot-carrier degraded performance.

4 Effect of gate length variation

4.1 Drain induced barrier lowering (DIBL)

Figure 8 shows the impact of gate length variation on DIBL for $V_{ds} = 0.5$ V and $V_{ds} = 0.8$ V. It is clearly evident from the figure that when the gate length reduces from 40 to 30 nm, reduction in DIBL is 1.8 % and when gate length reduces from 30 to 20 nm, the reduction in DIBL is 71.4 %. Thus TGRC-MOSFET shows significant reduction in DIBL as the gate length approaches sub 100 nm regime.

4.2 Electric field

Figure 9 shows that the variation of gate length from 40 to 30 nm enhances electric field at source side as compare to the drain side. A further reduction in gate length from 30 to 20 nm results in a significant enhancement in electric field. At the drain side, electric field is lower as compared

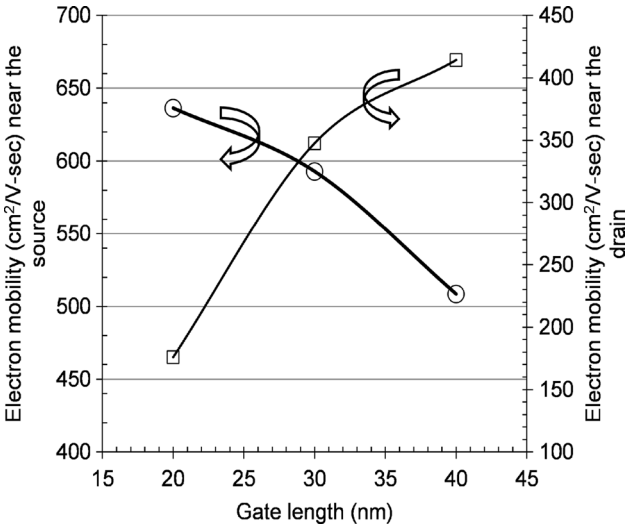


Fig. 11 Electric field variation along the channel for TGRC-MOSFET at different gate lengths

to source side when the gate length is reduced. The reduction in electric field at the drain end can be interpreted as a reduction in lower impact ionization, hot carrier effects, and higher breakdown voltage (Hajjiah and Huang 2004).

4.3 Electron velocity and electron mobility

Electron velocity of TGRC-MOSFET is observed with the gate length variation from 40 to 30 nm and from 30 to 20 nm as shown in Fig. 10. Electron velocity is higher at $L_G = 30$ nm as compared to $L_G = 40$ nm. Further reducing the gate length from 30 to 20 nm results an improvement in electron velocity at the source side as compared to drain side in TGRC-MOSFET (Gomez et al. 2009). Thus, enhancement in electron velocity at the source side makes the device faster hence; it is suitable for switching applications. Reduction of electron velocity at drain side interpreters less impact ionization hence, hot carriers reduced in TGRC-MOSFET.

Figure 11 shows the impact of gate length variation on electron mobility in TGRC-MOSFET. It is clearly evident from the figure that Electron mobility increases near the source and decreases near the drain with reducing the gate length. When gate length decreases from 40 to 30 nm then electron mobility enhances 16.7 % and a further reduction in the gate length from 30 to 20 nm increases the electron mobility increase by 7.3 % at the source side.

5 Conclusion

Here we focus on Transparent Gate architecture in which transparent gate is incorporated on conventional recessed chanel MOSFET for improved hot-carrier reliability of scaled MOS devices. Intensive 3-D TCAD device simulations have been performed to probe the internal transport conditions of CRC MOSFET and TGRC-MOSFET, and substantial interpretation is given to the internal behaviors observed in all the devices investigated. TCAD simulation reveals the deterioration in hot-carrier effects metrics such as electron velocity, potential, electron mobility and electron temperature. Further the gate length variation shows enhanced property when gate length reduces from 40 to 20 nm. This work presents TGRC-MOSFET a reliable device for high performance applications in CMOS technology.

Acknowledgments The authors would like to gratitude to Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University to carry out this work.

References

AI-Mistarihi MF, Rjoub A, AI-Taradeh NR (2013) Drain induced barrier lowering (DIBL) accurate model for nanoscale Si-MOSFET transistor. IEEE microelectronics (ICM), 2013 25th international conference. doi:10.1109/ICM.2013.6735011

AI-Mistarihi MF, Rjoub A, AI-Taradeh NR (2013) Drain induced barrier lowering (DIBL) accurate model for nanoscale Si-MOSFET transistor. IEEE, 978-1-4799-3570-3/13. doi:10.1109/ICM.2013.6735011

Arora N (2007) MOSFET models for VLSI circuit simulation theory and practice. Springer Verlag, Wien, New York, pp 87–89. ISBN 3-211-82395-6

Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008a) TCAD assessment of gate electrode workfunction engineered recessed channel (GEWE-RC) MOSFET and its multilayered gate architecture—part I: hot-carrier-reliability evaluation. IEEE Trans Electron Devices 55:2602–2612. doi:10.1109/TED.2008.2003085

Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008b) GEWE-RCMOSFET: a solution to CMOS technology for RFIC design based on the concept of intercept point. International conference on microwave, IEEE, pp 661–664. doi:10.1109/AMTA.2008.4762978

Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008) Solution to CMOS technology for high performance analog applications: GEWE-RCMOSFET, ACMD

Gautam R, Saxena M, Gupta RS, Gupta M (2013) Analytical model of double gate MOSFET for high sensitivity low power photosensor. J Semicond Technol Sci 13:500–510. doi:10.5573/JSTS.2013.13.5.500

Gomez L, Hashemi P, Hoyt JL (2009) Enhanced hole transport in short-channel strained-SiGe p-MOSFETs. IEEE Trans Electron Devices 56:2644–2651. doi:10.1109/TED.2009.2031043

Hajjiah AT, Huang AQ (2004) Novel SiC-Trench-MOSFET with reduced oxide electric field. IEEE, pp 340–344. doi:10.1109/PLASMA.2002.1030690

Krieger G, Sikora R, Cuevas PP, Misheloff MN (2001) Moderately doped NMOS (M-LDD)-hot electron and current drive optimization. IEEE Trans Electron Devices 38:121–127. doi:10.1109/16.65745

Long W, Ken KC (1997) Dual material gate field effect transistor (DMGFET). Electron devices meeting tech dig, pp 549–552. doi:10.1109/IEDM.1997.650445

Orouji AA, Kumar MJ (2005) Shielded channel double-gate MOSFET: a novel device for reliable nanoscale CMOS applications. IEEE Trans Device Mater Reliab 5:509–514. doi:10.1109/TDMR.2005.853505

Pammi SVN, Jung HJ, Yoon SG (2011) Low-temperature nanocluster deposition (NCD) for improvement of the structural, electrical, and optical properties of ITO thin films, pp 1059–1065. doi:10.1109/TNANO.2010.2103568

Pan Y, Ng KK, Wei CC (1994) Hot-carrier induced electron mobility and series resistance degradation in LDD NMOSFET's. IEEE Electron Device Lett 15:499–501. doi:10.1109/55.338416

Salehi A, Gholizadeh M (2011) Selectivity enhancement of MOSFET hydrogen sensor with transparent ITO gate. IEEE Sens J 11:1201–1205. doi:10.1109/JSEN.2010.2090947

Shee S, Bhattacharyya G, Dutta PK, Sarkar SK (2014) Threshold voltage roll-off and DIBL model for DMDG SON MOSFET: a quantum study. In: Proceeding of the 2014 IEEE students' technology symposium, pp 381–385. doi:10.1109/TechSym.2014.6808080

Shee S, Bhattacharyya G, Dutta PK, Sarkar SK (2014) Threshold voltage roll-off and DIBL model for DMDG SON MOSFET: a quantum study, pp 381–385. doi:10.1109/TechSym.2014.6808080

Thilakan P, Kumar J (1997) Studies on the preferred orientation changes and its influenced properties on ITO thin films. Vacuum 48:463–466

Torabi AJ, Joo M, Li X, Lim BS, Zhai L, Oentaryo RJ, Peen GO, Zurada JM (2013) A survey on artificial intelligence-based modeling techniques for high speed milling processes. IEEE Syst J, pp 1–12. doi:10.1109/JSYST.2013.2282479

ISSN: 2312-7694
Rupender et al, / International Journal of Computer and Communication System Engineering (IJCCSE), Vol. 2 (2), 2015, 356-359

Analytical Level Crossing Rate and Average Outage Duration for MRC Combiner in Log Normal Shadowed Channels

Rupender Singh
Department of Electronics & Communication
Delhi Technological University, Delhi, India
rupendersingh04cs39@gmail.com

Sanjay Kumar Soni
Department of Electronics & Communication
Delhi Technological University, Delhi, India
sanjoo.soni@gmail.com

***Abstract*— We provide an analytical approach for the evaluation of important second order statistical parameters, as level crossing rate (LCR) and average outage duration (AOD) of MRC combiner in Log-Normal shadowed channels. New expressions have been derived for LCR and AOD in closed form. The results are valid for an arbitrary number of independent identically distributed diversity branches in the presence of Log -Normal shadowed. The output of MRC combiner has been estimated using Fenton-Wilkinson method. LCR and AOD have been plotted and also tabulated for different number of diversity branches. The results that we provide in this paper are an important tool for measuring the performance of communication links in a log-normal shadowing.**

***Index Terms* – Level Crossing Rate (LCR), Average Outage Duration (AOD), Maximal Ratio Combining (MRC), Log-Normal shadowed (LN), Signal to Noise Ratio (SNR), Independent Identically Distributed (IID)**

I. INTRODUCTION

The log-normal distribution is often found to be the most suitable distribution to fit empirical fading channel measurements particularly for the indoor radio propagation environments [9]. The use of log-normal distribution [1], [10] to model shadowing which is random variable (SNR) can't be expressed in closed form solution for integrations involving in sum of random variables (SNRs) of MRC combiner ouput . This distribution (PDF) can be approximated by another log normal random variable using Fenton-Wilkinson method [3]. The level crossing rate (LCR) and average outage duration (AOD) are important metrics for evaluating the dynamic performance of diversity systems [11]. LCR and AOD have been considered in [11, 12-16] of different fading channels. These papers have considered Nakagami, Rayleigh, Rician, Rician-Log normal and Rayleigh-Log normal shadowed channels. A method to derive LCR and AOD for the log normal shadowed channel has been shown in [5]. Rice has given a method to find LCR and AOD of a fading channel using the joint PDF of the fading envelope and its time derivative [17]. But traditional Rician approach to estimate LCR and AOD is more complex and not suitable to characterize discrete channels [4, 6, 7].

In this work we have proposed a more reliable method for the statistical estimation of LCR and AOD. We have proposed a new and simple analytical approach to evaluate LCR and AOD of MRC combiner output in Log Normal (LN) shadowed channels. A simple joint CDF based approach for LCR and AOD has been shown. The MRC combiner output has been estimated using Fenton- Wilkinson method. Different values of Means and Variances have been tabulated for different number of diversity branches L. LCR and AOD has been plotted and tabulated for different number of diversity branches L.

II. SYSTEMS AND CHANNEL MODELS

Log-normal Distribution

A RV γ is log-normal, i.e. $\gamma \sim LN(\mu, \sigma^2)$, if and only if $\ln(\gamma) \sim N(\mu, \sigma^2)$. A log-normal RV has the PDF

$$p(\gamma) = \frac{\xi}{\sigma\gamma\sqrt{2\pi}} e^{-\frac{(10\log_{10}\gamma-\mu)^2}{2\sigma^2}} \qquad (1)$$

For any $\sigma^2 > 0$. The expected value of γ is

$$E(\gamma) = e^{(\mu+0.5\sigma^2)}$$

And the variance of γ is

$$Var(\gamma) = e^{(\sigma^2-1)} * e^{(2\mu+\sigma^2)}$$

Where ξ=10/ln 10=4.3429, μ(dB) is the mean of $10\log_{10}\gamma$, σ(dB) is standard deviation of $10\log_{10}\gamma$.

CDF of log-normal RV

$$P(\gamma) = Q\left(\frac{\mu - 10\log_{10}\gamma}{\sigma}\right) \qquad (2)$$

Where Q (.) is one dimentional standard Gaussian Q function.

III. MAXIMAL RATIO COMBINING

The total SNR at the output of the MRC combiner is given by:

$$\gamma_{MRC} = \sum_{i=1}^{L} \gamma_i \qquad (3)$$

$\gamma_{MRC}=\gamma 1+\gamma 2+\gamma 3+\gamma 4\ldots\ldots\ldots\ldots+\gamma L$.....(4)

Where L is number of IID branches.
Using Fenton-Wilkinson method, PDF of MRC combiner output is

$$p(\gamma_{MRC})=\frac{\xi}{\sigma_M\gamma_{MRC}\sqrt{2\pi}}e^{-\frac{(10log_{10}\gamma_{MRC}-\mu_M)^2}{2\sigma^2{}_M}} \quad (5)$$

And CDF of MRC combiner output is

$$P(\gamma)=Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right) \quad (6)$$

Table I μ_M and σ_M for different number of diversity branches [8]

Number of diversity branches L	μ_M	σ_M
2	*2.04*	0.85
4	*2.70*	0.65
6	*3.51*	0.55
8	*4.05*	0.48
10	*4.51*	0.44
15	*5.51*	0.36
20	*6.35*	0.31
25	*7.09*	0.28
30	*7.76*	0.26
50	*10.01*	0.20

IV. LEVEL CROSSING RATE AND AVERAGE OUTAGE DURATION

The AOD, $T_{\gamma th}$ (in seconds) is a measure of how long, on the average, the system remains in the outage state [1].

$$T_{\gamma th}=\frac{P_{out}}{N_{\gamma th}}$$

Where P_{out} is outage probability.
The envelope level crossing rate $N_{\gamma th}$ is defined as the expected rate (in crossings per second) at which the signal envelope crosses the level γ_{th} in the downward direction [2]

$$N_{\gamma th}=\int_0^\infty \dot{y}\, f_{\gamma,\dot{y}}(\gamma_{th},\dot{y})\, d\dot{y}$$

Where $f_{\gamma,\dot{y}}(\gamma_{th},\dot{y})$ is the joint PDF of γ and its time derivative.
Above definitions are Rician approach to estimate LCR and AOD. But this approach is not suitable for discrete channels. So for discrete sampled random process, LCR can be defined as the rate at which the envelope γ crosses a certain threshold γth in the positive or in the negative direction

$$N_y=\frac{P(\gamma 1\le \gamma th,\gamma 2\ge \gamma th)}{T} \quad (7)$$

Where $\gamma 1 \triangleq \gamma(t)$ and $\gamma 1 \triangleq \gamma(t+T)$ is second order statistics and T denotes sampling period.
Here $\gamma 1$ and $\gamma 2$ are uncorrelated IID random variables so their CDF is same as shown in (6).

Noting here

$$P(\gamma 1\le \gamma th,\gamma 2\ge \gamma th)=P(\gamma 1\le \gamma th)-P(\gamma 1\le \gamma th,\gamma th\ge \gamma 2) \quad (8)$$

Or

$$P(\gamma 1\le u,\gamma 2\ge u)=CDF\ of\ \gamma 1-Joint\ CDF\ of\ \gamma 1\ and\ \gamma 2$$

CDF of $\gamma 1$ for L diversity branches

$$P(\gamma 1\le \gamma th)=Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right) \quad (9)$$

And Joint CDF of $\gamma 1$ and $\gamma 2$ for L diversity branches

$$P(\gamma 1\le \gamma th,\gamma th\ge \gamma 2)=Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right) \quad (10)$$

Using (8), (9) and (10) in (7), LCR of MRC combiner output in log normal shadowed channel is

$$N_{\gamma th}=\frac{Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)-Q^2\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)}{T} \quad (11)$$

AOD is defined as

$$T_{\gamma th}=\frac{P(\gamma\le \gamma th)}{N_{\gamma th}} \quad (12)$$

Using (9) and (11) in (12) we have AOD of MRC combiner output in log normal shadowed channel is

$$T_{\gamma th}=\frac{T\,Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)}{\left(Q\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)-Q^2\left(\frac{\mu_M-10log_{10}\gamma_{MRC}}{\sigma_M}\right)\right)} \quad (13)$$

© IJCCSE All Rights Reserved

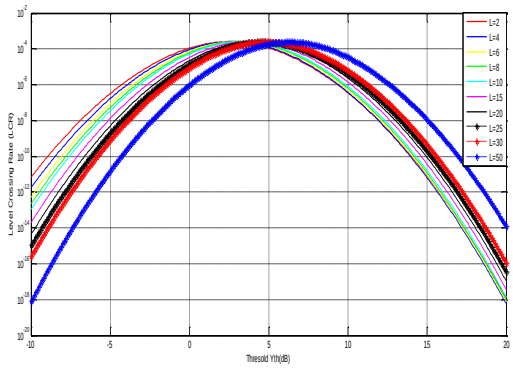


Fig 1. LCR of MRC combiner output in log normal fading channel T=1000sec

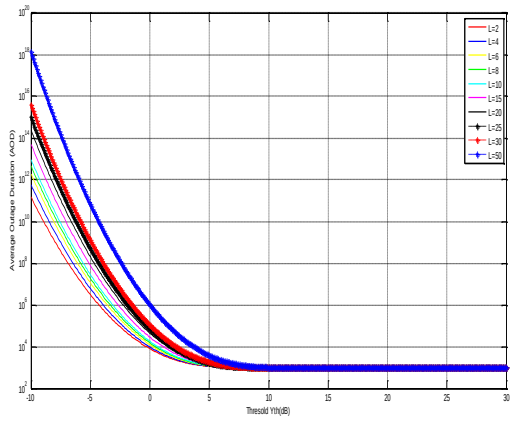


Fig . 2 AOD of MRC combiner output in log normal fading channel T=1000sec

V. CONCLUSION

This paper has established a process for estimating the LCR and AOD for MRC combiner output in log normal shadowed channel. Here we have used Table I which was given in [8]. LCR and AOD has been plotted and tabulated as well for different number of diversity branches L. For low threshold level, LCR decreases with increase in number of diversity branches L while for high threshold level (SNR), LCR increases with increase in number of diversity branches L. For low threshold level (SNR), AOD increases with increase in number of diversity branches L while for high threshold level, AOD decreases with increase in number of diversity branches L. So we can conclude that MRC combining improves system's performance in terms of LCR and AOD in log normal shadowed channel.

REFERENCES

[1] Marvin K. Simon, Mohamad Slim Alouni "Digital Communication over Fading Channels", Wiley InterScience Publication.
[2] A. Goldsmith, "Wireless Communications", Cambridge University Press, 2005.
[3] Neelesh B. Mehta, Andreas F. Molisch," Approximating a Sum of Random Variables with a Lognormal", IEEE Trans on wireless communication, vol. 6, No. 7, July 2007.
[4] W Jakes, Microwave mobile communications, IEEE Press classic reissue, Wiley, 1974.
[5] Cotton, S.L., Scanlon, W.G, "Higher Order Statistics for Lognormal Small-Scale Fading in Mobile Radio Channels"", Antennas and Wireless Propagation Letters, IEEE (Volume:6) , pp. 540 – 543, 2007.
[6] G L Stuber, "Principles of mobile communication", Norwell MA, Kluwer, 2001.
[7] M D Yacoub, J E V Bautista, and L Guerra de Rezende Guedes, "On higher order statistics of the Nakagami-m distribution," IEEE Trans. on Vehicular Technology, 48, May 1999.
[8] Rupender Singh, S.K. Soni, "Performance Analysis of MRC Combiner Output Signal in Log Normal Shadowed Fading", RAECE 2015 IIT Roorkee.
[9] Md. Zahurul I. Sarkar and T. Ratnarajah, "Secrecy Capacity Over Log-normal Fading Channel with Diversity Combining Techniques", 2013 IEEE Wireless Communications and Networking Conference (WCNC): PHY
[10] F. Hansen and F.I. Mano, "Mobile Fading-Rayleigh and Lognormal Superimposed", IEEE Trans. Vehic.Tech., Vol. 26, pp. 332–335, 1977.
[11] Daniel Benevides da Costa, Michel Daoud Yacoub, "Crossing Rates and Fade Durations for Diversity-Combining Schemes over α-μ Fading Channels", IEEE Trans Wireless Communications, vol. 6, no. 12, DECEMBER 2007.
[12] F. Graziosi, F. Santucci, "On SIR Fade Statistics in Rayleigh-Lognormal Channels", IEEE 2002.
[13] Cyril-Daniel Iskander, P. Takis Mathiopoulos, "Analytical Level Crossing Rates and Average Fade Durations for Diversity Techniques in Nakagami Fading Channels", IEEE Trans Wireless Communications, vol. 6, no. 12, DECEMBER 2007.
[14] Zoran Hadzi-Velkov, "Level Crossing Rate and Average Fade Duration of EGC Systems with Cochannel Interference in Rayleigh Fading", IEEE TCOM 2009.
[15] Norman C. Beaulieu, Xiaofei Dong, "Level Crossing Rate and Average Fade Duration of MRC and EGC Diversity in Ricean Fading", IEEE Trans Communications, vol 51 no 5, MAY 2003.

[16] Zoran Hadzi-Velkov, Nikola Zlatanov, George K. Karagiannidis, "Level Crossing Rate and Average Fade Duration of the Multihop Rayleigh Fading Channel", IEEE Communications Society ICC 2008 proceedings.

[17] S. Rice, "Statistical properties of a sine wave plus noise," Bell Syst. Tech. J., vol. 27, January 1948, pp. 109–157.

Appendix (Table II & Table III)

Table II Level Crossing Rate of MRC combiner output

$\gamma_{th} = -10dB$		$\gamma_{th} = -5dB$		$\gamma_{th} = -0dB$		$\gamma_{th} = 5dB$		$\gamma_{th} = 10dB$		$\gamma_{th} = 15dB$		$\gamma_{th} = 20dB$	
Number of diversity branches *L*	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	N_{yth}
2	6.707e-12	2	3.202e-07	2	1.099e-04	2	1.141e-04	2	3.531e-07	2	7.838e-12	2	1.179e-18
4	1.756e-12	4	1.679e-07	4	9.217e-05	4	1.247e-04	4	3.712e-07	4	6.24e-12	4	5.607e-19
6	5.328e-13	6	8.748e-07	6	7.348e-05	6	1.434e-04	6	5.058e-07	6	8.816e-12	6	7.295e-19
8	2.194e-13	8	5.354e-07	8	6.141e-05	8	1.574e-04	8	6.342e-07	8	1.146e-11	8	9.027e-19
10	1.071e-13	10	3.53e-08	10	5.202e-05	10	1.705e-04	10	8.01e-07	10	1.585e-11	10	1.304e-18
15	1.971e-14	15	1.29e-08	15	3.397e-05	15	2.006e-04	15	1.403e-06	15	3.597e-11	15	3.493e-18
20	4.27e-15	20	5.008e-09	20	2.2e-05	20	2.249e-04	20	2.394e-06	20	8.358e-11	20	1.045e-17
25	1.02e-15	25	2.004e-09	25	1.405e-05	25	2.417e-04	25	3.995e-06	25	1.97e-10	25	3.379e-17
30	2.526e-16	30	8.027e-10	30	8.783e-06	30	2.495e-04	30	6.505e-06	30	4.613e-10	30	1.119e-16
50	7.772e-19	50	1.539e-11	50	9.869e-07	50	1.912e-04	50	3.52e-05	50	1.141e-08	50	1.207e-14

Table III Average Outage Duration of MRC combiner output

$\gamma_{th} = -10dB$		$\gamma_{th} = -5dB$		$\gamma_{th} = 0dB$		$\gamma_{th} = 5dB$		$\gamma_{th} = 10dB$	
Number of diversity branches *L*	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}	**Number of diversity branches *L***	T_{yth}
2	1.491e+11	2	3.122e+06	2	7.954e+03	2	1.151e+03	2	1e+03
4	5.695e+11	4	5.955e+06	4	9.735e+03	4	1.171e+03	4	1e+03
6	1.877e+12	6	1.143e+07	6	1.252e+04	6	1.210e+03	6	1.001e+03
8	4.557e+12	8	1.868e+07	8	1.521e+04	8	1.243e+03	8	1.001e+03
10	9.334e+12	10	2.833e+07	10	1.816e+04	10	1.279e+03	10	1.001e+03
15	5.074e+13	15	7.754e+07	15	2.84e+04	15	1.385e+03	15	1.001e+03
20	2.342e+14	20	1.997e+08	20	4.443e+04	20	1.1519e+03	20	1.002e+03
25	9.8e+14	25	4.99e+08	25	7.017e+04	25	1.692e+03	25	1.004e+03
30	3.959e+15	30	1.246e+09	30	1.129e+05	30	1.916e+03	30	1.007e+03
50	1.287e+18	50	6.498e+10	50	1.012e+06	50	3.882e+03	50	1.038e+03

© IJCCSE All Rights Reserved

IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 17, Issue 5.Ver. I (May. 2015), PP 11-22
www.iosrjournals.org

Analytical study: Workplace Environment in IT Companies and Ranking of Companies using AHP –Topsis Method

Bhavna Saighal[1], M. K. Bhat[2] and Anjana Gupta[3]
[1] *(Research Scholar, Theerthankar Mahaveer University, Moradabad, India,)*
[2] *(Maharaja Agrasen Institute of Management Studies, GGSIP University Delhi, India)*
[3] *(Department of Mathematics, Delhi Technological University, Delhi, India)*

Abstract: *Workplace Environment that attracts, keep and motivate its Employees is considered the best one. Today's workplace is persistently changing. Sound workplace environment lead to employees' satisfaction which in turn helps to increase the productivity of the any Organization. The purpose of the present work is to study analytically the workplace environment in IT companies and its impacts on employee's performance and then ranking of the Companies. To analyse the workplace environment in IT sector various factors are involved like Physical, Economic, Behavioural and Motivational. This paper aims to explore the relationship between workplace environment and ranking of the top-10 listed IT companies in terms of Physical, Economic, Behavioural and Motivational factors using combined AHP-TOPSIS approach. Primary data were collected through the standard questionnaire from the employees of IT Companies. The presented hybrid multiple criteria decision making helps the employees to select the company effectively and objectively when multiple criteria are evaluated simultaneously. The first part of the projected methodology, i.e. AHP is constructive for determining the importance of each criterion and calculating weight of each criterion, while the second part with TOPSIS is useful for evaluating alternatives more precisely and help in ranking of the companies. The model can be implemented as an effective decision aid for potential candidate, investor and customers. Selecting a good company is a big challenge that should be evaluated very carefully.*

Keywords: *Analytical Hierarchy process(AHP), Employees Productivity, IT Sector, Office Environment, Technique of Order Preference by Similarity to Ideal Solution(TOPSIS), Workplace Design*

I. Introduction

Work is most important part of everyone's life. Work and the place of work call attention to many issues related to organizational psychology including job satisfaction, value of work life, human factors, working conditions, performance evaluation, motivation and leadership, and the physical and mental health of workers.

The management has to institute an environment that gives motivation to its employees for higher productivity. This liability lies with the Mangers at various levels of the organization. They have to think over and investigate the innovative working ways through which people can get pleasure from, feel like and have self-importance about what they do. The work environment affects employee self-esteem, efficiency and commitment- both definitely and disapprovingly and it may affect together physical and mental well-being. The present working surroundings is unlike, sole and frequently changing from time to time and context. The conservative employer and employee relationship of past days has been curved upside down. In growing economic there are so many work opportunities in front of employees, this has created a situation where employer needs the employees rather than employee needs employer. Hence, it becomes imperative to understand that work environment which is thought to be very important influence on employees working in IT Sector

The excellence and extent of job done by employees are influenced by the workplace environment. Deprived environmental conditions can cause incompetent worker efficiency as well as reduce their job satisfaction. An unsatisfied worker is a liability for any organization. He can never contribute up to his maximum for the development of the Organization. It may not be out of way to mention that a dissatisfy worker may spread negativity in the Organization which may in turn eclipse its productivity. The various factors influencing the workplace environment are Physical, Economic Motivational and Behavioural. All these factors are interlinked in such a way that no one of them can be treated separately. If these factors are congenial for the workers, the output of the Organization will definitely be at a higher level. The importance of these factors is become more important in case of organization involving creative thinking like IT Industry. The factors have an impact on the psychology of the workers.

IT Companies in India have played a vital role in Indian Economy. In the last 20 years of economic liberalization their contribution has been significant for the Indian Economy. This sector has helped the economy to maintain its growth rate at a higher level and now the IT companies have carved a place for

themselves in the economy.

IT Sector still bears a great unexplored potential for the economy. The easy availability of Technologists is making this sector quite lucrative for domestic as well as for foreign investor. It contributes nearly 60% of our GDP. The productivity in this sector can be further improved if the workers feel comfortable at the work place. The work place environment in a majority of IT offices is unsafe and unhealthy because of poorly planned workstations, inappropriate furniture, lack of ventilation, unsuitable lighting, excessive noise, insufficient safety measures in case of fire emergencies and lack of personal protective equipments. People working in such environment are prone to occupational diseases and it influences their performances. The productivity decreases due to this bad workplace environment. So the purpose of our research is to study carefully all these factor so that performance of employees working in IT sector can be enhanced and simultaneously give an in hand data for all the stakeholders.

II. Literature Review

Yi. Li., 2011 [1] pointed out that both companies and employees are increasingly concerned about the physical environment with the social development and continuous improvement. Empirical studies of enterprises in 12 provinces of China on the basis of the theoretical analysis were conducted. Correlation analysis was used to analysis of data and it showed that physical work environment has a significant impact on employee satisfaction. The workers contentment and performance also have a direct link, thus proving there was a significant correlation between the physical environment and performance of employees. Physical surroundings are the vital guarantee to uphold physical and mental health of employees, improve performance, and constitute the core competitiveness of enterprises.

Saleem, Ambreen. et. al., 2012 [2] conducted the study on the impact of physical environment on employees' productivity in higher education establishment of Khyber Pakhtoonkhawa (KPK) region of Pakistan. A common observation was that an improved workplace environment produces better results**.** The majority of the workplaces are planned according to the nature of the work. In corporate level, productivity is affected by many factors such as workers, technology, and health, safety moral and cultural aspects. Better workplace provides a better output. This study was conducted at individual level i.e., who directly engage in the education at university level. Performance/productivity was taken as dependent variable while furniture, room temperature, noise, lighting and other arrangements were taken as independent variables. Dependent and independent factors were calculated by the feedback from the educationists' of different Universities of KPK province of Pakistan, through the questionnaire. The respondents' replies were obtained by adding across all the item scores of the individual variables. The hypothesize associations among the study variables represented in the model which were tested using multiple regressions & factor analysis. Results reveal that there was a positive relationship between spatial arrangements and productivity. However, there is a pessimistic and considerable impact of noise and temperature on academicians 'productivity of higher education institutes of KPK province of Pakistan. Furniture and lighting has an irrelevant impact on employees' productivity, which shows that sample was not quite enough to explain this relationship significantly.

David McGuire and Lauren McLaren (2008) **[3]** examined the effect of the physical environment on employee commitment. It explores how favorable working conditions can affect an employee's sense of well-being which in turn can generate higher levels of employee commitment. A questionnaire instrument based upon previously validated measures was completed by 65 front line call centre employees. The Baron and Kenny four-step procedure for testing mediation effects was adopted. The statistical analysis confirms that employee well-being mediates the relationship between physical environment and employee commitment.

Jacqueline C. Vischer (2007) **[**4] Studied stress in the work environment and pay slight attention to factor of the physical environment in which work is done. Evidence accumulated that the physical environment of work affects both job performance and job contentment. This research reviewed theory and research bearing on stress in the workplace and explores how current theory might be applied to the relationship between worker behavior and physical features of the work environment. The research presented a hypothetical model of the worker–workspace relationship in which stress and comfort play a vital part, and propose a practical approach on which to base future empirical studies. Thus, by combining elements of work stress research with the environmental psychology of workspace, it was evident that a new area of study was opening up, namely the study of stress attributable to the design of workspace. As we find out more about how, when and why the buildings where people work affect their health and morale, so we would be able to help companies make more humane and cost-effective decisions about workspace.

Barry P. Haynes, (2008**)** [5] evaluated the impact that office comfort has on office occupiers' productivity. Author evaluates the literature that claims to make a linkage between the physical comfort of the office environment and the effect on the productivity of the office occupiers. Office comfort was initially discussed as a generic concept and subsequently broken down into sub-components. It was found that the evaluation of office comfort is a complex one. There appears to be no unanimously established explanation of

office comfort, and there is a clear need of conformity as to how office comfort should be measured. This research established that, there was sufficient data to support the claim that office comfort has an effect on productivity.

Workers in information technology (IT) - enabled services like business process outsourcing and call centers working with visual display units are facing health and psycho social disorders (C. Kesavachandran) 2006 [6], data from formerly published studies in journals and internet sources were examined to look at health problems and psycho-social problems among personnel working in IT-based services. In addition, author executed a questionnaire- based on pilot study. The accessible literature and the pilot study, together propose health disorders and psychosocial problems among workers of Business Process Outsourcing. It was found that Muscular-skeletal disorders, ocular disorders and psycho-social problems were the key health problems observed among IT professionals. The implementation of the course that include training of employees, concepts of ergonomics, health training and overcome the morbidity has become the need of today among workers in ITES

One way organizations increase their competitive advantage through innovative strategies that improve human performance Tim O. Peterson, Jon W. Beard, (2004) [7]. Human performance can be enhanced or constrained by situational factors that are introduced into the organization's work environment. Organization's workspace is one situational factor. The study examines the impact of a new workspace technology on individual privacy and on team interaction**.** It was found that the participants were generally satisfied with the visual privacy but not with the auditory privacy. The research originate that the participants were satisfied with the workspace's ability to facilitate team interaction.

Gretchen, M., Spreitzer. (1995) [8] developed and validated a multidimensional measure of psychological empowerment in the workplace. Second-order assenting factor analyses were conducted with two complementary samples to demonstrate the convergent and discriminated validity of four dimensions of empowerment and their contributions to an overall construct of psychological empowerment. Structural equations modelling were used to examine a homological network of psychological empowerment in the workplace. Tested hypothesis concerned key antecedents and consequences of the construct. Early support for the construct validity of psychological empowerment was found. Cronbah alphas and test-retest coefficient were used to assess the reliability of the empowerment measures. The research contributed to the literature by developing a conceptual definition of empowerment, measuring it, providing evidence of its construct validity and demonstrating its relationship to a number of antecedents and outcomes in the homological network.

Paul Roelofsen, 2002 [9] emphasized that one of the human need is to work in environment that allows people to perform their work optimally under comfortable conditions. Specified that structure and air conditioning systems are designed on the basis of a certain level of discomfort, this raises the key question 'What is the effect of the level of comfort on the productivity of people working in office environments?' He quantified this relationship as an aid to making choices regarding the working environment at strategic level within the facilities administration process, with meticulous emphasis on thermal conditions. It was found that the indoor environment has the biggest effect on productivity in relation to job stress and job dissatisfaction.

Alireza Bolhari et.al. (2012) [10] investigated occupational stress among information Technology (IT) professional and trying to measure the level of occupational stress and studying the relationship occupational stress level of IT professional and a) gender, b) age, c) work experience and d) stress management courses. The research was conducted among 236 IT professionals. The result suggested that stress reduction programs and strategies are inevitable due to high occupational stress levels. The association between strain level and gender, work experience, and stress management courses were approved through path analysis. Finally, implications of the study were discussed in terms of discussion and further research.

Nanjamari. K 2013 [11] emphasized on quality of the work environment because this is a place where workers spend mainly their time outside the home. Long hours are often spent in the working environment by IT employees. The organizations should help employees fight on-the-job-stress which can improve job satisfaction of its workers and thereby increases revenue for the company. There are other work circumstances that were used to determine job satisfaction, and these normally comprise how a company handle conflict, benefits, fair policies, level of interaction between management and employees, job safety, condition and accessibility of working tools and resources for performing work, growth and development opportunities in addition to flexibility, vacation, sick leave with pay, waged holidays, volunteer prospect, comp days, leave of absence, maternity leave, paternity leave, training and development. With all the factors mentioned, it is simple to observe why one's work situation seen as the most important factor of job satisfaction by many researchers. Consequently, study focused on investigating job satisfaction among information technology (IT) employees of Karnataka state. The primary objective of this study was to ascertain the levels of job satisfaction amongst information technology employees at an IT (information technology) and BPO (Business processing outsource) sector. The study undertaken to analyze the job satisfaction amongst information technology (IT) employees in Bangalore city in Karnataka state. The primary and secondary information was used to attain the objectives of

the study. The worldwide aspect of the study were based on the available articles and unpublished works on the theme etc. wholly in the study, books, journals and bulletins are considered to collect date. The primary date was collated through structured questionnaire and interview schedule in an unofficial environment to the IT professions.100 IT employees (1/2) of total employees were chosen randomly out of 200 IT employees in Bangalore city. The statistical tools like percentages and averages were used to analyze the data. The results of this study specify that general satisfaction is significantly associated with satisfaction levels about independence and fundamental factors.

III. Methodology

In this study workplace environment is categorized into four types: Physical, Economic, Behavioural, and Motivational. Physical environments consist of connectivity between users and workplace environments. Subcategories of physical environments include; Flexibility of reorganizing furniture, Comfortable furniture, Noise effect, Room Temperature, Ventilation, Wall Colour, Creach Facility, Recreational facilities, Study Material, Green Surroundings, Pantry facility, Waiting Room, Cleanliness and Distance between gadgets. Economic factors of the work environment directly affected the economic condition of the individual it includes; Compensation paid, Compensation reviewed, Adequate paid leaves, Gratuity policy, financial grants, Interest free loan, Insurance Schemes and Fringe benefits. Behavioural and Motivational environments are concerned with connectivity amongst users, colleagues and their organization. Subcategories of these environments includes; Working conditions, Reorganization as an individual, Suggestions of workers, Identification of Strength and Weakness, Employees as an asset, Innovative things at work , Praise and rewards, Work and personal life, Comfortable and open work Environment, Birthday and Anniversary remembrance and Support in personal crises. Behavioural Factors includes; Clear reporting structure, Support and team work, Company's objectives, Clear cut goals, Defining job responsibilities, Communication for improvement, Individual care and High achievement. We conducted three surveys. The first survey was general survey to find out the relevant factors for second survey. In second survey, consist of 9 point scale given by Saaty in 1980 [12] for analytical hierarchy process (AHP) and the third survey was conducted which was analysed by TOPSIS method. The respondents were requested to mark each response carefully to obtain convincing judgments.

3.1 Analytical Hierarchy Process (AHP)

To achieve our goal we are using technique of Analytical Hierarchy Process (AHP) given by Saaty in 1980 [12]. AHP present an absolute structure for structuring a decision problem to signify and quantify its elements. The result of AHP is a prioritized weighing of each decision alternative. The AHP convert this estimation to numerical values which can be processed and compared over the entire range of the problem. A numerical weight or priority is calculated for each element of the hierarchy which allow varied and repeatedly incommensurable elements to be compared with one another in a realistic and consistent way. The initial step in AHP is to formulate the problem as a hierarchy. The hierarchy comprises of an overall goal at the top level, alternatives for attaining the goal and criteria that relate the alternatives to the goal.

Once the hierarchy is formed, the participants make use of the AHP to establish priorities for all its nodes. For this, the elements of a problem are compared in pairs with respect o their relative impact on a property they share in common. The pair wise comparison is quantified in matrix form by using the scale of Relative Importance developed by Saaty [12] as shown in Table 1.

Table 1 Analytic Hierarchy Measurement Scale

Reciprocal measure of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance of one over another	Experience and judgment slightly favour one activity over another
5	Moderate importance	Experience and judgments moderately favour one activity over another.
7	Strong Importance	An activity is strongly favoured and its dominance is demonstrated in practice
9	Absolute Importance	The evidence favouring one activity over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values between two adjacent judgments.	When compromise is needed

During the elicitation process, a positive reciprocal matrix is formed in which the $(i,j)^{th}$ element a_{ij} is filled by the corresponding number from the Table 1 and the number is chosen according to the following criterion.

$$\begin{cases} a_{ij},\ if\ x_i\ dominates\ x_j \\ \frac{1}{a_{ij}},\ if\ x_j\ dominates\ x_i \\ 1,\ if\ x_i\ does\ not\ dominate\ x_j\ and\ x_j\ does\ not\ dominate\ x_i. \end{cases}$$

The matrix so formed is called the reciprocal matrix. It is used to compute the local priority weight of each criterion. The local priority weight (w) is the normalized eigen vector of the priority matrix corresponding to the maximum eigen value of the matrix. For detail reasoning of this account we refer to Ball and Noel [13], Bryson and Mobolurin [14], Lunging [15] and Saaty [12].

An attractive property of the priority matrix is that if in addition its elements are such that

$$a_{ij}\, a_{jk} = a_{ik}, \quad \forall\ i \le j \le k \qquad (1)$$

Then the derived priority vector w satisfies

$$w_i / w_j = a_{ij}, \quad \forall\ i < j \qquad (2)$$

Any reciprocal matrix satisfying "equation 1" is called consistent. However in practice, the priority matrix seldom satisfies "equation 1", thereby making it more important to define some relaxed measuring of consistency check. Saaty introduced the concept of consistency index CI of a reciprocal matrix as the ratio $\frac{\lambda_{max} - n}{n-1}$, Where λ_{max} and 'n' respectively stand for the maximum eigen value and order of the reciprocal matrix.

Table 2 Random Consistency Index (RI)

N	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

3.1.1 Development of Hierarchical Structure

Planning a hierarchy means of modeling the result graphically at hand. It consists of an overall goal, alternatives to achieve the goal, and a collection of factors or criteria that relate the alternatives to the goal.

The above identified factors are now considered as the relevant criteria and sub criteria and are used to formulate an appropriate AHP model for selecting the factors of workplace environment of IT Companies. Theoretically, all the factors can be included in the AHP-based model, as the AHP methodology will enable us to compare and prioritize them. However, it is not practical to include all factors as they increase the number of pair wise comparisons and the related computational effort. It is also possible that assessment biases may occur in obtaining the pair wise comparison judgments from evaluators. Through AHP application we will come to the weights/importance of each factor in the set of hierarchy. The study includes Employees of IT sectors of all levels. The ranking shows the selection of the most preferred Work Environment where the candidate will prefer to work in.

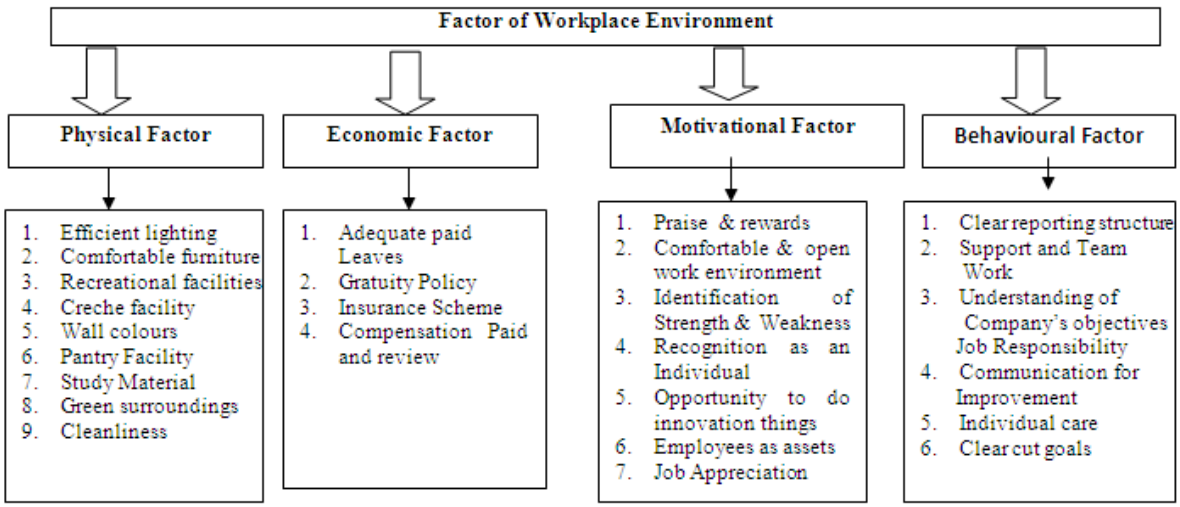


Figure 1 Hierarchy for Workplace Environments and Employees Satisfaction in I.T. Sector

To evaluate the hierarchy (Fig. 1), surveys were conducted to rate each attribute to others in a series of pair wise comparison using a Scale from 1 to 9 (Table 1).An approach of Analytic Hierarch process is applied to find out the weights of each criterion a different levels of hierarchy. Survey results were analyzed for each level of the hierarchy and reciprocal matrices were generated. An aggregated reciprocal matrix was developed for each set of matrices having (C.R<0.2) by taking geometric mean of corresponding values for further calculations. The local priority weights were derived from a series of pair wise comparisons involving all the nodes.

Table 3 Relative Importance of Criteria at Main Level

Attributes	Physical	Economic	Motivational	Behavioural	Priority
Physical	1	1.2268	1.993	0.8247	0.280096
Economic	0.81512	1	2.6502	0.7363	0.265692
Motivational	0.50175	0.37733	1	0.432	0.126308
Behavioural	1.21256	1.35814	2.31481	1	0.327904

λmax = 4.02791 C.I.=0.00930169 C.R. = 0.010341

Table 4, 5, 6 and 7 gives the corresponding matrix at level 2 and weights are also calculated at Physical, Economic, Motivational and Behavioural level. Further, the Global weights of the criteria are calculated as given in table 8 and the graphic presentation of all the weights are shown in "Fig. 2".

Table 4 Relative Importance of Criteria at Physical Factor

Attributes	Efficient lighting	Comfortable furniture	Recreational facilities	Creche facility	Wall colours	Pantry Facility	Study Material	Green surroundings	Cleanliness	Priority
EL	1	2.7737	2.6982	3.06921	2.674	2.4777	2.577	1.7814	2.2747	0.232777
CF	0.360529	1	1.9732	1.4529	0.899	1.1268	1.569	0.57988	0.7562	0.0999892
RF	0.370617	0.50678	1	0.82795	0.715	0.7015	0.672	1.06943	0.4558	0.0692839
CF	0.325817	0.68827	1.2078	1	0.696	1.2648	1.355	0.634639	0.4904	0.0797482
WC	0.373972	1.11175	1.3982	1.4366	1	0.6635	1.193	0.81686	0.4594	0.0883255
PF	0.4036	0.88740	1.4253	0.790591	1.507	1	0.783	0.666111	0.5609	0.0853747
SM	0.387919	0.637137	1.4869	0.737647	0.837	1.2770	1	0.90815	0.5622	0.0832981
GS	0.561355	1.72449	0.9350	1.5757	1.2242	1.5012	1.101	1	0.5614	0.110378
C	0.439618	1.3224	2.1937	2.03882	2.1765	1.7826	1.778	1.78113	1	0.150825

λmax =9.21042 C.I.=0.0263028 CR= 0.018139

Table 5 Relative Importance of Criteria at Economic Factors

Attributes	Adequate paid leaves	Gratuity Policy	Insurance Scheme	Compensation Paid & review	Priority
Adequate paid leaves	1	1.6921	1.42208	1.21252	0.31749
Gratuity Policy	0.590982	1	1.60058	0.668706	0.215905
Insurance Scheme	0.703195	0.624775	1	0.527314	0.169164
Compensation Paid & review	0.824731	1.49543	1.8964	1	0.297441

λmax =4.04435 C.I.=0.014784 C. R. = 0.01642

Table 6 Relative Importance of Criteria at Motivational Factors

Attributes	Praise & rewards	Comfortable & open work envi.	Identification of Strength & Weaknesses	Recognition as an Individual	Opportunity to do innovation things	Employees as an assets	Job Appreciation	Priority
P&R	1	1.9705	2.1118	2.3812	1.584	2.1369	1.578	0.240385
C&OWE	0.5074	1	2.21972	2.2755	1.032	1.2528	1.539	0.172863
ISW	0.4735	0.4505	1	1.9198	0.690	0.793371	1	0.110631
RI	0.4199	0.4394	0.520888	1	0.693	0.71797	0.663	0.083106
OIT	0.6310	0.9683	1.44728	1.4422	1	0.894403	1.164	0.140798
EA	0.4679	0.7982	1.26044	1.3928	1.118	1	0.958	0.128292
JA	0.6335	0.6497	1	1.5071	0.858	1.0432	1	0.123924

λmax =7.09454 C.I.=0.0157559 C. R. = 0.01193

Table 7 Relative Importance of Criteria at Behavioural Factors

Attributes	Clear reporting structure	Support and team work	Understanding of Company's Objective	Job Responsibility	Communication for improvement	Individual care	Clear cut goals	Local Weights
CRS	1	2.22358	2.29879	1.83491	1.83713	2.15125	1.894	0.24965
STW	0.449725	1	2.23325	2.11573	0.96783	1.63783	1.56337	0.17313
UCO	0.435011	0.447778	1	2.16387	0.730607	0.84733	0.96404	0.11274
JR	0.544986	0.472651	0.462136	1	0.872064	0.698827	0.63964	0.08880
CI	0.544328	1.03324	1.36872	1.1467	1	0.918529	1.14221	0.13318
IC	0.464846	0.610564	1.18018	1.43097	1.0887	1	0.693648	0.11580
CCG	0.527983	0.639644	1.0373	1.56338	0.875495	1.44165	1	0.12668

λmax= 7.17197 C.I. = 0.0286614 C.R. =0.02171

Similar matrices are worked out for all the criteria at level 3. Also global priority weights for each criterion are calculated by multiplying its local weight by its upper level criterion's global priority weight.

The evaluation process finally generates the global weights for each requisite criterion of interest as shown in Table 9 and the graphic presentation of all the weights are shown in "Fig. 2".

Table 9: Global weights of the Criteria

Ranking	Attributes	Local Weights	Global Weights
1	Adequate paid Leaves	0.31749	0.084355
2	Clear reporting structure	0.24965	0.08186
3	Compensation Paid and review	0.29744	0.079027
4	Efficient lighting	0.23277	0.065201
5	Gratuity Policy	0.21590	0.057364
6	Support and team work	0.17313	0.05677
7	Insurance Scheme	0.16916	0.044945
8	Communication for Improvement	0.13318	0.04367
9	Cleanliness	0.15082	0.042246
10	Clear cut goals	0.12668	0.04154
11	Individual care	0.11580	0.037973
12	Understanding of Company's Objective	0.11274	0.036968
13	Green surroundings	0.11037	0.030917
14	Praise and rewards	0.24038	0.030363
15	Job Responsibility	0.08880	0.029119
16	Comfortable furniture	0.09998	0.028007
17	Wall colours	0.08832	0.02474
18	Pantry Facility	0.08537	0.023913
19	Study Material	0.08329	0.023332
20	Creche facility	0.07974	0.022337
21	Comfortable and open work environment	0.17286	0.021834
22	Recreational facilities	0.06928	0.019406
23	Opportunity to do innovation things	0.14079	0.017784
24	Employees as assets	0.12829	0.016205
25	Job Appreciation	0.12392	0.015653
26	Identification of Strength and Weaknesses	0.11063	0.013974
27	Recognition as an Individual	0.08310	0.010497

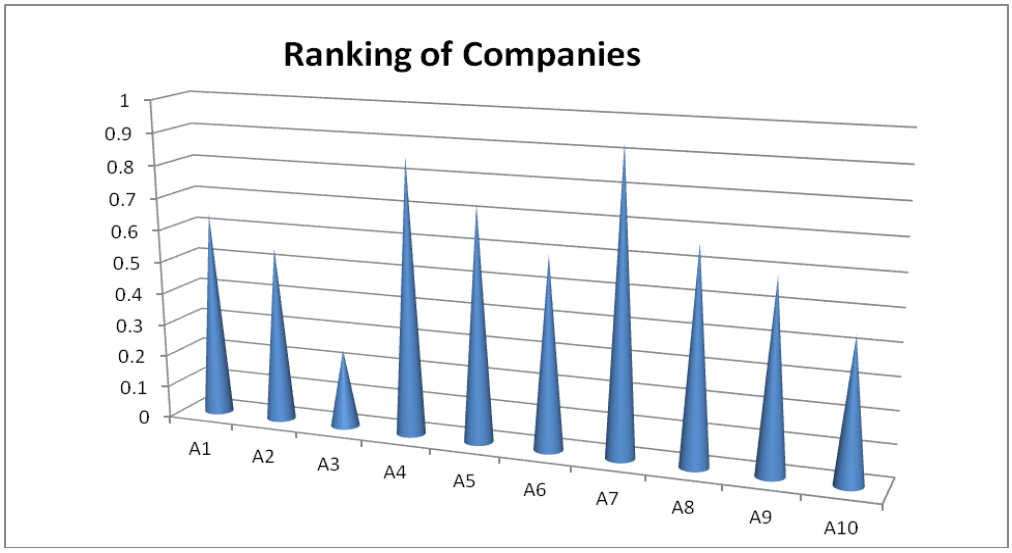


Figure 2

On the basis of above calculations it can be held that factors like adequate paid leave, clear reporting structure, compensation paid and reviewed have a dominant role to play for the employees satisfaction and thereby have a clear cut impact on their productivity compared to the factors like recognition as an Individual, identification of strength and weakness and job appreciation have least impact on the employees satisfaction and thereby on their productivity. It seems that the employees in the IT sector are giving more preference to economic consideration than Motivational Behavioural and Physical Factors.

In the next section, on the basis of factors deprived from AHP , we will rank the companies by using TOPSIS Method that will result in ranking of the companies will give in hand data to the employees, customer and investors.

3.2 TOPSIS (Technique for Order Preference by Similarity to the Ideal Solution)

TOPSIS (Technique for Order Preference by Similarity to an Ideal Solution) is a popular approach to AHP and has been widely used in the literature. According to Shih et al. [11], some advantages of TOPSIS are as follows:

- A sound logic that embodies the rational of human choice
- A simple computation process that can be easily programmed into a spreadsheet
- A scalar value that accounts for both the best and worst alternative at the same time.

Let A_i, i = (1,..., m) be a set of m alternatives, C_j, j = (1, ..., n) be n evaluation criteria and W_j, j = (1, ..., n) be the associated criteria weights as shown in "Table 10". It is also assumed that the weights W_j of the criteria have been determined by an appropriate method (this is not a part of the TOPSIS methodology). Furthermore x_{ij} is the performance measure of alternative A_i with respect to criterion C_j

Table 10 : Decision Matrix

(A_i) ↓	C_1	C_j	C_n
	W_1	W_j	W_n
A_1	x_{11}		x_{1n}
A_i		x_{ij}	
A_m	x_{m1}		x_{mn}

The procedural steps of TOPSIS are summarized as follows:

Step 1: Obtain an evaluation x_{ij} as crisp score for each alternative corresponding to the respective criterion. Here x_{ij} is the performance measure of i^{th} alternative with respect to j^{th} criterion.

Step 2: Normalize the decision matrix using the following equation:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_i x^2_{ij}}} \qquad (i = 1,..., m), \quad (j = 1,..., n) \qquad (1)$$

Normalization of the decision matrix transforms the various attribute dimensions into non dimensional attributes, which allows comparison across the attributes.

Step 3: Form the weighted normalized decision matrix using the formula

$$v_{ij} = w_j r_{ij} \quad (i=1,..., m), \quad (j=1,..., n) \qquad (2)$$

Step 4: Determine the positive ideal and negative-ideal solutions.

$$A^+ = [(\max_i (v_{ij}): j \in J), (\min_i (v_{ij}): j \in J'), (i=1,..., m)]$$

$$= [v_1^+, v_2^+, ..., v_j^+, ..., v_n^+] \qquad (3)$$

$$A^- = [(\min_i (v_{ij}): j \in J), (\max_i (v_{ij}): j \in J'), (i=1,..., m)]$$

$$= [v_1^-, v_2^-, ..., v_j^-, ..., v_n^-] \qquad (4)$$

- J be the set of benefit attributes or criteria (more is better).
- J´ be the set of negative attributes or criteria (less is better)

Step 5: Calculate the separation measures for each alternative.

Ideal Separation

$$S_i^+ = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^+)^2}, \quad (i=1,..., m) \qquad (5)$$

Negative Ideal Separation

$$S_i^- = \sqrt{\sum_{j=1}^{n} (v_{ij} - v_j^-)^2}, \quad (i=1,..., m) \qquad (6)$$

Step 6: Calculate the relative closeness to the ideal solution.

$$C_i^* = \frac{S_i^-}{S_i^+ + S_i^-}, \quad (i=1,..., m) \qquad (7)$$

Step 7: A set of alternatives can now be preference ranked according to the descending order of C_i^*.

Application of TOPSIS method

We asked 10 experts to give their opinion on 5 point scale for 10 companies based on the given 27 criteria discussed in the previous section. Now we have 10 matrix of each expert which is aggregated by Arithmetic Mean. The resultant aggregated matrix is given in "Table 11" and corresponding normalized matrix (by using equitation 1) given in "Table 12". Weighted normalized Decision Matrix is calculated by using "equation 2" is given in "Table 13". It is to note here that the global weights calculated through AHP in section 3.1 have been used to calculate the weighted normalized decision matrix. Determine the positive ideal and negative-ideal solutions by using "equations 3 & 4" given in "Table 14 and 15" respectively. Finally the ranking of the alternatives is given in the last "Table-16" and also "Fig. 3" depicts the final result via graphic representation.

Table 11 Average Mean of the Data

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27
A1	4	3.2	4	4	4	4	3	2	4	4	4	4	3	5	3	3	2	3	3	3	3	4	3	4	4	3	3
A2	4	3	3	4	1	3	2	3	4	2	3	4	3	3	4	3	3	3	4	2	3	3	3	2	3	3	3
A3	1	2	3	4	1	3	2	2	4	2	4	4	3	2	2	2	3	3	3	3	2	2	3	3	2	2	2
A4	3.9	3.9	4.1	3.6	4.6	3.8	4.5	3.5	4.8	3.6	3.9	3.7	2.9	3.9	3.8	3.8	3.7	4.1	4.7	3.9	4.4	4.2	3.8	3.9	4.1	3.8	4.1
A5	1	2	2	3	1	2.2	1	2.2	3.4	3.4	2.5	1.9	2.1	3.4	2.5	1.1	1.8	2.2	3.1	1.3	2.2	1	2.2	3.4	3.4	2.5	1.9
A6	4.3	2.1	3.8	2.9	3.4	3.1	3.4	3.3	3.9	3.1	4.3	2.1	3.8	2.9	3.4	3.1	3.4	3.3	3.9	3.1	3.1	3.8	2.9	3.4	3.1	3.4	4
A7	4.3	2.1	3.8	2.9	3.4	3.1	3.4	3.3	3.9	3.1	4.3	2.1	3.8	2.9	3.4	3.1	3.4	3.3	3.9	3.1	3.1	3.8	2.9	3.4	3.1	3.4	4
A8	4.3	4	4.1	3.6	5	4.3	4	4.1	4.4	5	4.3	4	4.1	3.6	5	5.3	2.3	3.2	3.3	2.8	4.5	4	4	3.8	3.9	3.8	3.9
A9	4	3.8	4	3.5	5	4.5	4.3	3.1	3.4	3.3	3.9	3.1	4.3	2.1	3.8	2.9	3.4	3.1	3.4	3.3	3.9	3.1	3.8	3.9	3.8	3.9	5.3
A10	1.7	2.7	2.5	4.3	1.2	3.7	4	3.4	4	3.9	3.4	4	4	3.4	4	3.9	3.4	4	2.1	2.8	2.5	2.6	1.7	2.1	2.8	2.5	2.6

Table-12 Normalized decision Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27
A1	0.36	0.34	0.36	0.35	0.37	0.36	0.29	0.21	0.32	0.37	0.33	0.37	0.27	0.48	0.27	0.29	0.21	0.29	0.27	0.33	0.29	0.38	0.31	0.38	0.37	0.30	0.27
A2	0.36	0.32	0.27	0.35	0.09	0.27	0.19	0.31	0.32	0.18	0.25	0.37	0.27	0.29	0.35	0.29	0.32	0.29	0.36	0.22	0.29	0.29	0.31	0.19	0.28	0.30	0.27
A3	0.09	0.21	0.27	0.35	0.09	0.27	0.19	0.21	0.32	0.18	0.33	0.37	0.27	0.19	0.18	0.19	0.32	0.29	0.27	0.33	0.19	0.19	0.31	0.28	0.19	0.20	0.18
A4	0.35	0.41	0.37	0.32	0.43	0.34	0.43	0.36	0.38	0.33	0.32	0.34	0.26	0.37	0.34	0.36	0.39	0.40	0.42	0.42	0.43	0.40	0.39	0.37	0.38	0.38	0.37
A5	0.09	0.21	0.18	0.26	0.09	0.20	0.10	0.23	0.27	0.31	0.21	0.18	0.19	0.32	0.22	0.11	0.19	0.21	0.28	0.14	0.21	0.10	0.22	0.32	0.32	0.25	0.17
A6	0.39	0.22	0.35	0.25	0.32	0.28	0.32	0.34	0.31	0.28	0.36	0.20	0.35	0.28	0.30	0.30	0.36	0.32	0.35	0.34	0.30	0.36	0.30	0.32	0.29	0.34	0.36
A7	0.39	0.22	0.35	0.25	0.32	0.28	0.32	0.34	0.31	0.28	0.36	0.20	0.35	0.28	0.30	0.30	0.36	0.32	0.35	0.34	0.30	0.36	0.30	0.32	0.29	0.34	0.36
A8	0.39	0.43	0.37	0.32	0.47	0.39	0.38	0.42	0.35	0.46	0.36	0.37	0.37	0.34	0.44	0.51	0.24	0.31	0.30	0.30	0.43	0.38	0.41	0.36	0.37	0.38	0.35
A9	0.36	0.40	0.36	0.31	0.47	0.41	0.41	0.32	0.27	0.30	0.32	0.29	0.39	0.20	0.34	0.28	0.36	0.30	0.31	0.36	0.38	0.30	0.39	0.37	0.36	0.39	0.48
A10	0.15	0.29	0.23	0.38	0.11	0.33	0.38	0.35	0.32	0.36	0.28	0.37	0.37	0.32	0.35	0.37	0.36	0.39	0.19	0.30	0.24	0.25	0.17	0.20	0.26	0.25	0.23

Table-13 Weighted Normalized Decision Matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27
Weights taken from AHP	*0.31*	*0.24*	*0.29*	*0.23*	*0.21*	*0.17*	*0.16*	*0.13*	*0.15*	*0.12*	*0.11*	*0.11*	*0.11*	*0.24*	*0.08*	*0.09*	*0.08*	*0.08*	*0.08*	*0.07*	*0.17*	*0.06*	*0.14*	*0.12*	*0.12*	*0.11*	*0.08*
A1	0.11	0.08	0.11	0.08	0.08	0.06	0.05	0.03	0.05	0.04	0.04	0.04	0.03	0.11	0.02	0.03	0.02	0.02	0.02	0.02	0.05	0.02	0.04	0.05	0.00	0.03	0.02
A2	0.11	0.08	0.08	0.08	0.02	0.05	0.03	0.04	0.05	0.02	0.03	0.04	0.03	0.07	0.03	0.03	0.03	0.02	0.03	0.02	0.05	0.02	0.04	0.02	0.00	0.03	0.02
A3	0.03	0.05	0.08	0.08	0.02	0.05	0.03	0.03	0.05	0.02	0.04	0.04	0.03	0.05	0.01	0.02	0.03	0.02	0.02	0.02	0.03	0.01	0.04	0.03	0.00	0.02	0.01
A4	0.11	0.10	0.11	0.07	0.09	0.06	0.07	0.05	0.06	0.04	0.04	0.04	0.03	0.09	0.03	0.03	0.03	0.03	0.03	0.03	0.07	0.02	0.05	0.04	0.00	0.04	0.03
A5	0.03	0.05	0.05	0.06	0.02	0.03	0.02	0.03	0.04	0.04	0.02	0.02	0.02	0.08	0.02	0.01	0.02	0.02	0.02	0.01	0.04	0.01	0.03	0.04	0.00	0.03	0.01
A6	0.12	0.05	0.10	0.06	0.07	0.05	0.05	0.04	0.05	0.03	0.04	0.02	0.04	0.07	0.02	0.03	0.03	0.03	0.03	0.02	0.05	0.02	0.04	0.04	0.00	0.04	0.03
A7	0.12	0.05	0.10	0.06	0.07	0.05	0.05	0.04	0.05	0.03	0.04	0.02	0.04	0.07	0.02	0.03	0.03	0.03	0.03	0.02	0.05	0.02	0.04	0.04	0.00	0.04	0.03
A8	0.12	0.10	0.11	0.07	0.10	0.07	0.06	0.06	0.05	0.06	0.04	0.04	0.04	0.08	0.04	0.05	0.02	0.02	0.02	0.02	0.07	0.02	0.06	0.04	0.00	0.04	0.03
A9	0.11	0.10	0.11	0.07	0.10	0.07	0.07	0.04	0.04	0.04	0.04	0.03	0.04	0.05	0.03	0.03	0.03	0.02	0.02	0.03	0.06	0.02	0.05	0.04	0.00	0.04	0.04
A10	0.05	0.07	0.07	0.09	0.02	0.06	0.06	0.05	0.05	0.04	0.03	0.04	0.04	0.08	0.03	0.03	0.03	0.03	0.02	0.02	0.04	0.01	0.02	0.02	0.00	0.03	0.02

Table-14 The positive ideal solution

A1	0.01
A2	0.01
A3	0.02
A4	0.01
A5	0.03
A6	0.01
A7	0.01
A8	0.01
A9	0.01
A10	0.01

Table-15 The negative-ideal solution

A1	0.02
A2	0.01
A3	0.00
A4	0.03
A5	0.00
A6	0.02
A7	0.02
A8	0.03
A9	0.03
A10	0.01

Table -16- Ranking of Companies

Ranking	Alternative	Relative Closeness to the ideal solution of 10 alternative is computed using equation 5 & 6
1	A7	0.93641827
2	A4	0.85759309
3	A5	0.73196203
4	A8	0.66539337
5	A1	0.63673304
6	A6	0.59559667
7	A9	0.58803685
8	A2	0.54935141
9	A10	0.44149284
10	A3	0.24358419

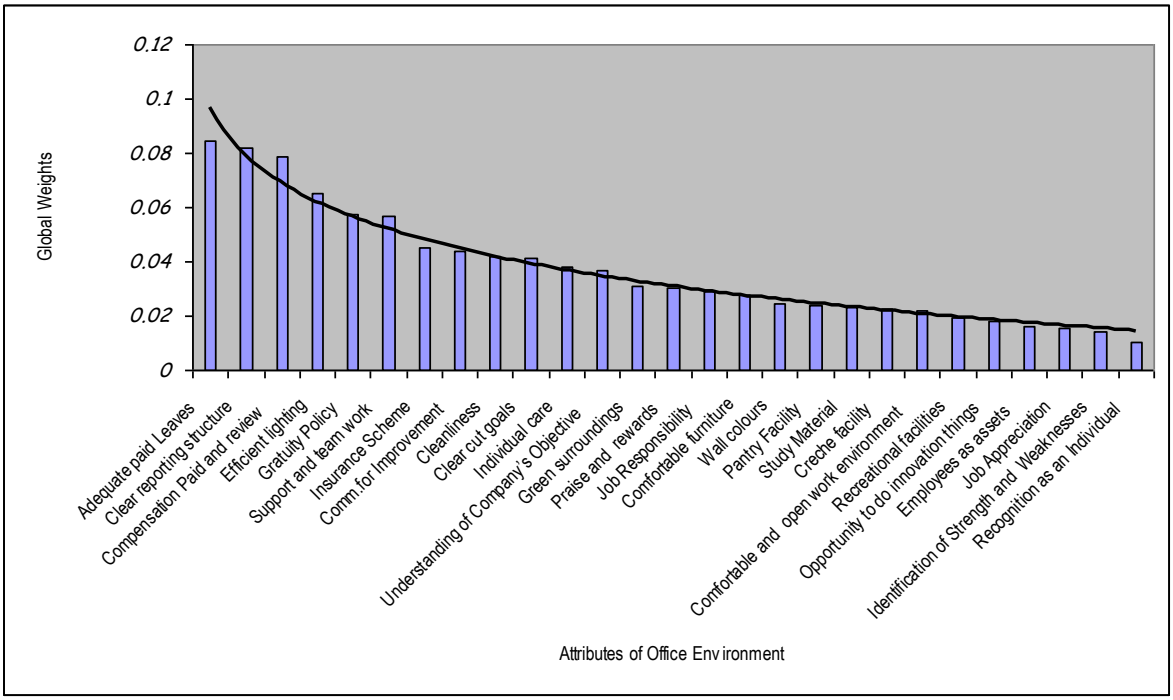


Figure 3

IV. Conclusion

In this paper we have suggested the way of weighing and ranking of companies. The aim of the current research is to develop the decision model for prospective candidate, investor and customers. In the current paper Analytic Hierarchy Process (AHP) and the Technique for Order Preference by Similarity to Ideal Solution method is used for ranking 10 IT companies. By applying the AHP the decision maker can examine the strengths and weakness of company by comparing them with respect to appropriate criteria and sub criteria. The model enables us to visualize the impact of various criteria on the final ranking and determines the level of importance of each criterion based on survey analysis. Ranking of companies make a distinction the distress from the healthy ones. Ranking of the Companies promotes Competition and market effectiveness. By ranking of companies Candidates will have enough indication which company to join and making his/her career there. Investors and market-operator will get guild as to where to put money for better return. Customer will also get good information as to with whom they should enter in the contact and develop their software. To support the calculation at advance level we have used TOPSIS method.

Reference

[1]. Yi. Li., Empirical study of the impact of physical environment on the employees performance, in proceeding of Emergency Management and Management Sciences (ICEMMS), 2011 2nd IEEE International Conference, August, 8-11, 2011, Beijing, China Available : IEEE Xplore, http://www.ieee.org.

[2]. Ambreen Saleem et al., Impact of Internal Physical Environment on Academicians Productivity in Pakistan, European Journal of Business and Management, 4(2), 2012, 43-62.

[3]. McGuire David and McLaren Lauren, The impact of physical environment on employee commitment in call centres: The mediating role of employee well-being, Team Performance Management: An International Journal, 15, (1/2), 2009, 35 – 48.

[4]. Jacqueline C. Vischer, The effects of the physical environment on job performance: towards theoretical model of workspace stress, Stress and Health, 23 (3), 2007, 175-184.

[5]. Barry P. Haynes, An Evaluation of the Impact of the Office Environment on Productivity, Facilities, 26 (5/6), 2008, 178 – 195
[6]. C. Kesavachandran et al., Working Conditions and Health among Employees at Information technology--enabled services: a review of current evidence, Indian Journal of Medical Sciences, 60 (7), 2006, 300-30.
[7]. Tim O. Peterson and Jon W. Beard, Workspace Technology's impact on individual privacy and team interaction, Team Performance Management: An International Journal, 10(7/8), 2004, 163 – 172.
[8]. Gretchen, M., Spreitzer, Psychological empowerment in the workplace: Dimensions, measurements and validation, Academy of Management Journal, 38(5), 1995, 1442 – 1465.
[9]. Paul Roelofsen, Performance Loss in Open-Plan Offices due to Noise by Speech, Journal of Facilities Management, 6(3),2008,202 – 211.
[10]. Alireza Bolhari et al., Occupational stress Level among Information Technology Professionals in Iran, International Journals of Information and Electronics Engineering, l 2(5), 2012.
[11]. H.S. Shih., H. J. Syur and E.S. Lee, An Extension of TOPSIS for group decision making, Mathematical and Computer Modeling, 2007, 45(7-8).
[12]. T. L. Saaty, the Analytic Hierarchy Process (McGraw Hill, 1980).
[13]. V.C. Ball, J. Noel and Srinivasan, Using the analytic hierarchy process in house selection, Journal of Real Estate Finance and Economics, 9, 1994, 69-85.
[14]. N. Bryson and A. Mobolurin A, An approach to using the analytic hierarchy process for solving multiple criteria decision making problems, European Journal of Operational Research, 76, 1994, 440-454.
[15]. F. Lunging, Analytical hierarchy in transportation problems, an application for Istanbul, Urban Trans- portation Congress of Istanbul, 2,1992,16-18.

Wireless Pers Commun
DOI 10.1007/s11277-015-2733-3

Ant Colony Optimization Based Orthogonal Directional Proactive–Reactive Routing Protocol for Wireless Sensor Networks

Aarti Jain[1] · **B. V. Ramana Reddy**[2]

© Springer Science+Business Media New York 2015

Abstract Routing protocols for wireless sensor networks are important in addressing the various quality-of-service (QoS) issues pertaining to different applications. The most important QoS issues while designing routing protocols for WSN are energy awareness, scalability and network lifetime. However to deal with these issues the solutions provided in related literature have certain inherent disadvantages like high control overhead, low packet delivery ratio and requirement of global location information. In order to resolve these issues, we propose an orthogonal transmission based scalable, lightweight and energy aware routing protocol named as OD-PPRP which does not require global location information and has low control overhead. The proposed protocol OD-PRRP has the characteristics of both reactive and proactive routing protocols and utilizes fuzzy logic and Ant Colony Optimization to identify energy efficient and optimal paths. The simulation results show in both static and dynamic environment, OD-PRRP has better network lifetime, low end to end transmission delay, less overhead and high packet delivery ratio than other state of art QoS aware routing protocol viz. EARQ, EAODV and EEABR.

Keywords Ant colony optimization · Energy efficient routing · Fuzzy logistics · Optimal hop length · Orthogonal transmission · Wireless sensor networks

✉ Aarti Jain
rtjain2001@gmail.com

B. V. Ramana Reddy
bvrreddy64@rediffmail.com

[1] Department of ECE, Ambedkar Institute of Advanced Communication Technologies and Research, Delhi, India

[2] USET, USICT, Guru Gobind Singh Indraprastha University, Delhi, India

Published online: 22 May 2015

1 Introduction

A wireless sensor network comprises of a distributed collection of wireless sensor nodes which have limited sensing, processing and communicating capabilities. The main task of sensor nodes is to sense, collect and process information and then relay the processed information to the base station. This communication is performed in a single or multi hop fashion depending on the location of the sensing nodes as well as of base station. Sensor nodes in WSN are battery powered and in most of cases, it is not possible to recharge or change batteries. Thus energy consumption is the most significant parameter, which needs special attention while designing communication protocols for WSNs. Moreover, in wireless sensor networks thousands of sensors are randomly deployed for a given application. With this high density of nodes, communicated data is likely to suffer communication failures due to contentions in shared communication medium. Besides these, changes in network topology due to node destruction or battery depletion require exchange of state information for maintaining and updating the links in network [1–3]. All these characteristics of wireless sensor networks make energy efficiency, scalability and optimal resource management as main research challenges while designing protocols for network layer.

To address the above mentioned challenges, existing routing protocols usually characterize the network with metrics such as residual energy, reliability, communication cost or delay [4]. For finding the routes that are energy efficient, reliable and short, a lot of control messages are broadcasted and exchanged between nodes. In most of the existing protocols it is assumed that the location information of the entire network is known to all the sensor nodes. This location information is assumed to be obtained either from the installed GPS receiver on nodes or from any other localization method. However, these existing solutions are costly in implementation and have accessibility and scalability related issues. Thus, there is a need for a routing methodology that provides an acceptable path which is energy efficient and does not require explicit positioning and high control overhead.

Considering the above, we have proposed a new routing method that avoids excessive overhead in maintaining complete link state information. The proposed method does not require global position information and ascertains efficient and balanced use of energy resources. The method named Orthogonal Directional Proactive Reactive Routing Protocol (OD-PRRP), optimally utilizes the characteristics of both the proactive and reactive routing model. In order to relax information requirements such as co-ordinate space embedding and node localization we have chosen the nodes to transmit packets in orthogonal directions. The orthogonal transmission leads to less congestion and contention losses, which in turn makes the system more scalable. Also, orthogonal transmissions achieve connectivity between source and destination with high probability even in sparse networks having voids and require only $O(N^{3/2})$ total state information to maintain an N-node network [5].

In OD-PRRP, routing information is proactively maintained by certain selected "*proactive*" nodes using Ant colony optimization. These proactive nodes are periodically selected by using random back-off scheme [6]. Subsequently, the balance nodes (with no route information) reactively identify the route with the help of routes already identified by these selected proactive nodes. The overhead are controlled as only few selected "proactive" nodes are utilized for proactive route maintenance. This responsibility of being *proactive* nodes is periodically shared among the nodes in the network to avoid non uniform energy consumptions. For the nodes other than the selected proactive nodes, the

delays associated with reactive methods are addressed at the time of route identification as any node having route information itself sends the route reply to the respective node on behalf of the sink node. Thus the nodes other than the proactive nodes utilize the methodology prescribed in AODV [7] method. However the delays associated with AODV method are addressed by proactive route maintenance by few selected proactive nodes as discussed above.

Further in OD-PRRP, Ant colony optimization [8], which is a widely accepted optimization method, has been applied to choose between different available routes. In wireless sensor networks multiple parameters i.e. energy efficiency, path length and hop length are to be considered while making routing decisions. The multiple parameters increase the complexity. To address this, we have modeled our network as directed fuzzy graph. In the proposed fuzzy graph the same is addressed by representing edge weight as a function of both residual energy as well as hop distance.

The performance of OD-PRRP has been analyzed by comparing the performance with state of art routing methods such as EARQ [9], AODV [10] and EEABR [11] by using MATLAB simulations. The simulation results indicate that the proposed method results in enhanced network lifetime, low end to end transmission delay, high delivery ratio, low overhead and balanced load distribution. Moreover, the performance of the proposed method is evaluated in dynamic environment i.e. after incorporating node mobility. The simulation done in dynamic environment indicate that the performance of the proposed protocol is reasonably well in terms of network lifetime, end to end transmission delay, control overhead and packet delivery ratio.

The main contributions of this paper are as follows:

1. Less memory requirement, low congestion and contention losses and no requirement of global location information have been ensured by using orthogonal transmission.
2. The advantages of both the proactive routing and reactive routing methodology i.e. fast response and low overhead respectively have been harnessed by combining the use of both these methodologies.
3. Ant colony optimization is used to find a path which is energy efficient and does not leads to excessive delay in terms of hops.
4. Modeling of network as Fuzzy Graph has led to efficient modeling and leads to less complexity.
5. The performance of OD-PRRP has been analyzed by comparing the performance with state of art routing methods EARQ and EAODV by using MATLAB simulations in both static and dynamic environment i.e. after incorporating node mobility.

This paper is organized as follows. Section 2 presents related work. Section 3 presents the system model. In Sect. 4, network modeling and detailed description of the proposed method OD-PRRP is given. In Sect. 5 simulation results are presented and effectiveness of our work is evaluated. Conclusions are given in Sect. 6.

2 Related work

Routing protocols are responsible for discovering and maintaining the routes in the network. Most of the routing protocols available in literature are either reactive or proactive in nature and use Omni directional communication mode. *Proactive routing* protocols are those protocols in which each node stores the routing information to every node in the network in its storage tables and keeps on refreshing these tables periodically. Some

examples of *proactive routing* protocols are Destination-Sequenced Distance Vector (DSDV) [12], Optimized Link State Routing (OLSR) [13] and Topology Broadcast Reverse Path Forwarding (TBRPF) [14]. *Reactive or on-demand routing* protocols are those protocols in which routing information is acquired by a node at the time, when it actually has some data to communicate. Some examples of *reactive routing* protocols are such as Dynamic Source Routing (DSR) [15], Adaptive On-Demand Distance Vector (AODV) [7], On-Demand Multicast Routing Protocol (ODMRP) [16] and Cluster Based Routing Protocol (CBRP) [17]. In addition to these basic routing protocols, in literature Hybrid routing protocols are also proposed. These Hybrid routing protocols divides the network into clusters, trees or zones and proactively or reactively maintain routes at different stages. GPSR (Greedy perimeter stateless routing) [18], ZRP (Zone Routing Protocol) [19] and Directed diffusion [20] are the examples of hybrid routing protocols.

In wireless sensor networks, due to limited available energy resources and inherent QoS requirements, researchers have developed various QoS and energy aware routing protocols. In most of these methods existing proactive and reactive routing methods are modified to make them energy efficient and QoS aware. EARQ [9] is a kind of proactive routing. It provides real-time, reliable delivery of a packet, while considering energy awareness. In EARQ, a node estimates the energy cost, delay and reliability of a path to the sink node, based only on information from neighboring nodes. Then, it calculates the probability of selecting a path, using the estimates. To achieve real-time delivery, only paths that may deliver a packet in time are selected. To achieve reliability, a source node may send a redundant packet via an alternate path. In [10] Li et al. has presented an enhanced AODV (EAODV) routing protocol for wireless sensor networks. In EAODV method the proper transmission energy required for communication between two nodes is calculated using dimidiate search arithmetic. Multiple routes between source and sink are found on demand using AODV protocol and finally a cost effective route is selected using accumulative summing method. The delay associated with finding the route is high as final path selection executes the iterative process dimidiate search arithmetic for selecting proper transmission power and then selects the route on the basis of accumulative power required for transmission. In [21] SAR algorithm is proposed to minimize the average weighted QoS metric throughout the lifetime of the network. It takes into consideration both energy resources, QoS metric on each path and priority level of each packet. However, the protocol suffers from the overhead of maintaining the tables and states at each sensor node and hence not suitable for densely deployed WSNs. Maximum lifetime energy routing proposed by Chang and Tassiulas [22] is a network flow base routing approach. The main objective of the approach is to maximize the network lifetime. To maximize network lifetime a new link cost as a function of node remaining energy and the required transmission energy is formulated. In [23] multiple-objective fuzzy decision making based routing protocol (MOFDMR) has been proposed for WSNs. In MOFDMR nodes periodically executing HEED [24] protocol for cluster heads selection and then multiple trees rooted from these cluster heads are established. Out of multiple paths, MOFDMRW chooses a path through complete comparisons of energy reserves, time latency and the reliability using multiple-objective fuzzy decision making and location information. However, due to periodic re-clustering, establishment of new routes for every new round may increase the overhead and can lead to inefficient utilization of energy. In [25], the author has proposed two online (reactive) fuzzy based routing algorithms. First algorithm FML attempts to maximize the WSN life time while the second algorithm FMO is for optimizing lifetime as well as energy conservation. Due to purely reactive nature of the protocol Delay and control overhead is very high and hence not suitable for delay sensitive applications. Jabbar et al.

[26] has proposed REAR, to achieve the real-time routing with energy awareness. It follows the customized proactive routing with in-network processing and localized behavior. Actual node position is required in this method. Time, Hop Count and Energy are taken as selection parameters with priorities in the respective order. Due to proactive nature and concept of mobility, overhead in maintain a route is very high. In [27], in order to enhance network life time, the author has proposed a centralized reactive routing method that uses fuzzy logic and A-star algorithm. Optimal path is determined using residual energy of nodes, number of hops between source and destination and current traffic load on nodes The base station determine the routing schedule based on the information received from nodes. But, due to centralized approach, control overhead in this solution is very high.

In paper [28], Network rOle-based Routing Algorithm (NORA) and Network role-based Routing Intelligent Algorithm (NORIA) are developed to perform role assignment during route establishment and maintenance. NORA performs decision making by using a series of mathematical operations whereas NORIA, follows a fuzzy-rule-based paradigm. It is a kind of proactive routing protocol and different types of control messages are broadcasted during route finding process. FEAR [29] protocol is a fuzzy based energy as well delay sensitive reactive protocol. FEAR consumes much less bandwidth or have less control overhead as compare to proactive protocols, but the delay in determining a route is substantially large. Routing decisions in FEAR are based on the fuzzy distance to the Base station in terms of hops as well as remaining energy of nodes on the path towards the base station.

In literature, some ant colony optimization based routing methods have been proposed for WSNs. The main objective of these methods is to optimize energy and delay metric. In [11], an energy efficient Ant colony optimization based routing method (EEABR) has been proposed. The method is proactive in nature and periodically each node sends a colony of artificial ants through the WSN to look for paths between the sensor nodes and a sink node. The selection of path by ant is based upon path length and energy efficiency. The EEABR protocol builds a routing tree with optimized energy branches through various iterations. However, proactive nature and route maintenance by all nodes results in the very high overhead. In [30], a centralized chain-based routing protocol, (AntChain) has been proposed for WSNs. In this algorithm, the base station finds routes of all the sensor nodes by using ACO. After forming the routes, the information is broadcasted to respective sensor nodes. Being a centralized approach, the method is not suitable for large scale WSNs as it results in high overhead. Moreover, the routing paths are formed by using only the energy metric and without considering the length of the path which leads to high end-to-end delay. In [31], an ant colony optimization based method ("E&D ANTS") has been proposed with primary objective of minimizing the delay in delivering a fixed number of data packets. In order to optimize energy and delay, the reinforcement learning (RL) algorithm has been introduced to train the model. The key idea of E&D ANT's method is the consideration of both energy and delay in wireless communication network to update the nodes' pheromones. In [32], an ant routing algorithm based on deflection angle is proposed for WSNs. In this algorithm, the energy, distance, and deflection angle are considered for the selection of routing path. However, due to proactive route maintenance, the routing overhead is very high in [31, 32] algorithms.

Cheng et al. [5] have introduced an Orthogonal Rendezvous Routing Protocol (ORRP) for meshed wireless networks. ORRP is a lightweight-but-scalable routing protocol utilizing directional communications to relax information requirements such as coordinate space embedding and node localization. The ORRP source and ORRP destination send route discovery and route dissemination packets respectively in locally-chosen orthogonal

directions. Connectivity happens when these paths intersect (i.e. rendezvous). However, in ORRP energy awareness and various QoS provisions are not considered as protocol is developed for simple mesh networks. Our proposed protocol uses orthogonal directional communication to increase delivery ratio by reducing contention losses, increase system throughput by maximizing spatial reuse and decrease energy consumption. Moreover, our proposed method has characteristics of both proactive and reactive routing characteristics and route finding is by using ACO which lead to minimization of both delay and excessive control overhead and results into energy and delay efficient paths.

3 System Assumptions

We consider a WSN with a large number of sensor nodes randomly deployed over a 2-D planer area. All the sensor nodes are identical in computational capabilities and starts with equal initial energy σ, $\sigma > 0$. The sensor nodes can communicate with each other through short-range radios. After initialization of network, each node $u \in V$discovers its respective one hop neighbor nodes N_u. Transmission between node u and neighboring node v is possible only if current energy of transmitting node $ce(u)$ is greater or equal to energy required for transmitting one unit of data. In this paper we use log-distance path loss model presented in [33], which is one of the most widely used models in network simulations and theoretical analysis. According to this model, power received by a remote node at distance d meters from the sender can be expressed as follows:

$$P_r(d) = P_o \times \left(\frac{d_o}{d}\right)^{\alpha} \tag{1}$$

where, P_o is the received signal power at the distance d_o from a transmitter and α is the path loss exponent which varies from 2 to 6.

In addition to above following assumptions have been made:

1. The received signal strength is measurable and hence nodes can adjust their transmission power.
2. Nodes transmit and receive packets in orthogonal directions with directional antenna.
3. All sensor nodes are location-unaware.
4. Every node in the network knows (1) its 1-hop neighbors and (2) direction/interface to send packets so that it reaches the respective neighbor.
5. Each node has its own local perception of direction with respect to a local North.

4 Proposed Method

In this section we discuss the network modeling and the proposed routing method OD-PRRP in detail.

4.1 Modeling of Network

We have modeled the WSN as directed Fuzzy vertex and edge graph [34] $\tilde{G}(V, E, \mu(E))$, where V represents the set of n nodes and $E \subseteq V \times V$ is the set of edges. The main characteristic of Fuzzy vertex and edge graph is that it considers both vertex set and edge

set as fuzzy sets and allows the attachment of fuzzy membership function (value) to both vertices and edges. The resulting degree of an edge of the said fuzzy graph is the function of its own primary edge degree and the degree of the two vertices connected by the incremented edge.

In this paper, we have considered directed fuzzy vertex and edge graph for which $\mu_{\tilde{G}}(E_{uv})$ represents the resulting degree of an edge between two nodes u and v. $\mu_{\tilde{G}}(E_{uv})$ is a function of its own primary edge degree $\mu_{\tilde{E}}(e_{uv})$ and the degree of the end node connected by the incremented edge; $\mu_{\tilde{V}}(v)$. We define $\mu_{\tilde{V}}(v)$ as the membership degree of a node v in a fuzzy set $\tilde{V}$ of nodes having *high* residual energy. Similarly $\mu_{\tilde{E}}(e)$ is the membership degree of an edge e in a fuzzy set $\tilde{E}$ of edges having *desirable* edge length (optimal transmission power required to transmit data). Here *high* and *desirable* represents the linguistic variables [35] for defining fuzzy node set $\tilde{V}$ and edge set $\tilde{E}$ respectively. In the next subsection we present the definition of fuzzy set and membership function and then one by one define the fuzzy membership values assigned to nodes i.e. $\mu_{\tilde{V}}(v)$, edges i.e. $\mu_{\tilde{E}}(e)$ and $\mu_{\tilde{G}}(E)$ respectively.

4.1.1 Fuzzy set and membership function

Fuzzy set theory provides mathematical tools for carrying out approximate reasoning processes, when available information is uncertain, incomplete or imprecise. These features of fuzzy set theory can add more flexibility and capability in modeling wireless sensor networks where most of the information regarding network is collected from a dynamic topology. The mathematical frame work to describe Fuzzy Logic was suggested by Zadeh [35] in his seminal paper entitled "Fuzzy Sets" (Zadeh 1965). According, if X is a collection of objects denoted by x, a fuzzy set $\tilde{A}$ in X can be defined as a set of ordered pairs:

$$\tilde{A} = \left\{ \left(x, \mu_{\tilde{A}}(x) \right) | x \in X \right\} \tag{2}$$

where $\mu_{\tilde{A}}(x)$ is called the membership function of x in $\tilde{A}$ that maps X to the membership space M. The range of the membership function is a subset of the non-negative real numbers whose supremum is finite. The said membership function is not limited to values between 0 and 1 but if $sup_x\left(\mu_{\tilde{A}}(x)\right) = 1$, the fuzzy set $\tilde{A}$ is called normal i.e. $\mu_{\tilde{A}} : X \rightarrow [0, 1]$ (including 0, 1). In comparison to crisp set, fuzzy set is a 'vague boundary set' [36].

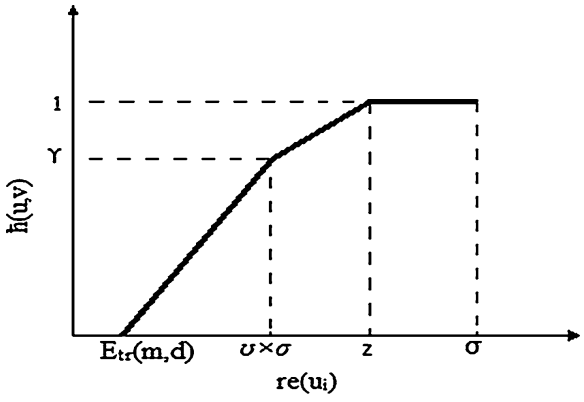


Fig. 1 Membership function for node set with high energy

4.1.2 Fuzzy membership function of node $\mu_{\tilde{V}}(v)$

In the proposed method a fuzzy set $(\tilde{V})$ (nodes with high residual energy) is defined on vertex set V and a fuzzy membership function $\mu_{\tilde{V}}(v)$ (Fig. 1) is assigned to nodes $v \in V$ corresponding to its current residual energy $re(v)$. A high membership value is assigned to a node having a large amount of residual energy. At the time of network initialization, the residual energy of all the nodes is high and is near to σ, hence the membership function value 1 is assigned to these nodes till their residual energy decreases to a level say z. After level z, the membership value is first made to fall at a low rate. Subsequently, when the residual energy level decreases below another threshold point $\mho \times \sigma$ (where $\mho \in [0, 1]$ is an algorithmic parameter) [25], the corresponding membership is decreased at a sharper rate. This behavior of the membership function is chosen to strongly discourage the selection of nodes for special responsibilities that have depleted their energy beyond a certain threshold value. The threshold point may be altered by adjusting the value of $\mho$. The characteristic Equations showing the membership degree assigned to a node v are depicted in Eq. 3.

$$\mu_{\tilde{V}}(v) = \begin{cases} 1 & : \quad if \;\; z < re(v) \leq \sigma \\ 1 - \left(\dfrac{1-\gamma}{1-\mho}\right) * \left(1 - \dfrac{re(v)}{\sigma}\right) & : \quad if \;\; \mho * \sigma < re(v) \leq z \\ \dfrac{\gamma}{\mho * \sigma - E_{tr}(m,d)} * (re(v) - E_{tr}(m,d)) & : \quad if \;\; E_{tr}(m,d) < re(v) \leq \mho * \sigma \\ 0 & : \quad if \;\; re(v) \leq E_{tr}(m,d) \end{cases} \tag{3}$$

Here,

$$re(v) = ce(v) - E_{tr}(m,d) \tag{4}$$

where $ce(v)$ denote current energy and $re(v)$ denotes predicted residual energy after one unit of data transmission by node v respectively, and $\theta \in [0, 1]$, $\gamma \in [0, 1]$ are algorithmic parameters. $E_{tr}(m,d)$ is the average energy consumed in one unit of data transmission.

4.1.3 Fuzzy Membership function of an edge $\mu_{\tilde{E}}(e)$

Wireless Sensor networks consist of nodes with limited energy resources. Sensor nodes dissipate energy while sensing, receiving, sending and processing data. The most energy-consuming component is radio module while the energy requirement for computation is substantially low. The energy consumption while transmitting 1 bit of data on the wireless channel is equivalent to energy required to execute thousands of cycles of CPU instructions [1, 2]. Thus the energy consumption for transmission is the determining factor for network life. With the advancement in VLSI technology, all modern radio transceivers have capability of adjusting the transmission power [37]. Thus data transmission towards destination can be done either with high transmission power (small number of larger hops) or low transmission power (large number of smaller hops). However, both high power transmission and low power transmission methods have their inherent advantages and disadvantages. Routing with short hops minimizes the transmission energy but overall energy consumption (transmission + reception) increases with the communication distance. On the other hand, transmitting packets using long hops reduces the reception cost (due to decrease in number of nodes involved in data routing), but it requires high transmission power and also leads to more contention losses, congestion losses and

attenuation losses. By optimization of hop lengths, energy can be saved which in turn extends the network lifetime [38]. For calculating optimal hop length we have used Eq. (1) i.e. log-distance path loss model to define the minimum power required to communicate over a given distance between two nodes. According to this model in order to guarantee the signal detection at the receiver node, the transmitted signal by node A must be at least equal to the value such that received power by node B is at least equal to sensitivity threshold P_t of the receiver or we can write:

$$P_t = P_x \times \left(\frac{d_o}{x}\right)^{\alpha} \tag{5}$$

Here, P_x denotes the minimum transmitted energy requited to transmit up to distance x between sender and receiver and is given by:

$$P_x = P_t \times \left(\frac{x}{d_o}\right)^{\alpha} \tag{6}$$

Though energy cost for receiving data is low as compared to transmission cost, but is not negligible. The energy cost for receiving data can be represented equivalent to transmitting over a distant r [39]. Thus reception energy cost P_R can be represented as:

$$P_R = P_t \times \left(\frac{r}{d_o}\right)^{\alpha} \tag{7}$$

Figure 2 represents two different topologies of multi-hop routing between two distant nodes, a source A and a destination B, with a distance d. The first topology uses n *short* hops of length x to transmit data from A to B; While the second uses m *long* hops of length y i.e. $n = \lambda m$ and $y = \lambda x$, so we can write:

$$P_t = P_x \times \left(\frac{d_o}{x}\right)^{\alpha} = P_y \times \left(\frac{d_o}{y}\right)^{\alpha} \quad for \quad 2 \leq \alpha \leq 6 \tag{8}$$

Thus, transmission power required to communicate at distance x and y respectively is given by:

$$\left.\begin{aligned} P_x &= P_t \times \left(\frac{x}{d_o}\right)^{\alpha} \\ P_y &= P_t \times \left(\frac{y}{d_o}\right)^{\alpha} \end{aligned}\right\} \tag{9}$$

Since $y = \lambda x$ using Eq. (9) we can write:

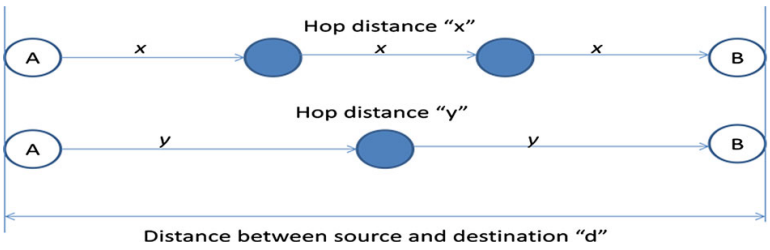


Fig. 2 Communication between node A and node B using long hop and short hop

$$P_y = P_x \times \lambda^{\alpha} \tag{10}$$

Using Eqs. (6) and (7), the energy required to communicate data from node A to node B using n short hops of length x is given by:

$$P_{short} = nP_x + nP_R \tag{11}$$

This can be written as

$$P_{short} = \lambda m \times P_x + \lambda m \times P_R \tag{12}$$

Using Eq. (10), this can be written as:

$$P_{short} = m \times P_y \left[\frac{1}{\lambda^{\alpha-1}} + \lambda \frac{P_R}{P_x}\right] \tag{13}$$

Similarly, power required to transmit data from A to B using m *long* hops is given by

$$P_{long} = mP_y + mP_R \tag{14}$$

This can be written as

$$P_{long} = m \times P_y \left[1 + \frac{P_R}{P_y}\right] \tag{15}$$

Now we have calculated the difference in power requirements using n small hops and m long hops

$$P_{short} - P_{long} = m \times P_y \left[\frac{1}{\lambda^{\alpha-1}} + \lambda \frac{P_R}{P_y}\right] - m \times P_y \left[1 + \frac{P_R}{P_y}\right] \tag{16}$$

This result into

$$P_{short} - P_{long} = m \times P_y \left[\frac{1}{\lambda^{\alpha-1}} - 1 + \frac{P_R}{P_y}\right] \tag{17}$$

Since $m \times P_y > 0$, therefore, $P_{short} - P_{long} > 0$, if and only if

$$\frac{1}{\lambda^{\alpha-1}} - 1 + \frac{P_R}{P_y} > 0 \tag{18}$$

Therefore, $\frac{P_R}{P_y} > 1 - \frac{1}{\lambda^{\alpha-1}}$

Using Eqs. (16), (17), (18) we get

$$y < \frac{r}{\sqrt[\alpha]{1 - \frac{1}{\lambda^{\alpha-1}}}} \tag{19}$$

Since $y = \lambda x$, Eq. (19) results into:

$$x < \frac{r}{\lambda \times \sqrt[\alpha]{1 - \frac{1}{\lambda^{\alpha-1}}}} \tag{20}$$

Since $y > x$, results obtained in Eqs. (19) and (20) propose that optimal energy consumption can be achieved when the hops lengths are included in the interval given by Eq. (21). Results indicate that optimal hop length is dependent on attenuation constant and energy consumed in receiving data.

$$\left]\frac{r}{\lambda \times \sqrt[\alpha]{1-\frac{1}{\lambda^{\alpha-1}}}}, \frac{r}{\sqrt[\alpha]{1-\frac{1}{\lambda^{\alpha-1}}}}\right[\tag{21}$$

In randomly deployed wireless sensor networks, network may have voids or low density of nodes at some places and it may not be tenable to have fixed transmission range based upon the optimum hop distance value. In order to address this limitation resulting from random deployment and for energy efficiency (optimal and required hop length), in OD-PRRP transmission power has been divided into different levels i.e., from PTXMIN to PTXMAX. Thus, a node can transmit within this range of minimum and maximum transmission power, so as to maintain connectivity within a randomly deployed network. Additionally, by choosing the next hop node at optimal hop distance the system becomes more adjustable and energy efficient for randomly deployed network.

In the proposed method, fuzzy logistics has been used to formulate this situation. Assuming that it is required to have optimal transmission power equal to or near to P_{Tx}^{opt}, where P_{Tx}^{opt} lies in range

$$\left[\frac{r}{\lambda \times \sqrt[\alpha]{1-\frac{1}{\lambda^{\alpha-1}}}}, \frac{r}{\sqrt[\alpha]{1-\frac{1}{\lambda^{\alpha-1}}}}\right].$$

The function depicted in Eq. (22) below presents a possible value of fuzzy membership function $\mu_{\tilde{E}}(e)$ in fuzzy set of edges $\tilde{E}$ corresponding to optimal transmission power. It is shown that the highest membership function value of $\mu_{\tilde{E}}(e) = 1$ is for $P_{Tx} = P_{Tx}^{opt}$. Transmission range is allowed up to minimum P_{Tx}^{Min} and maximum P_{Tx}^{Max}, but with a dropping membership function.

$$\mu_{\tilde{E}}(e) = \begin{Bmatrix} 0 & P_{Tx} < P_{Tx}^{\min} \\ \dfrac{P_{Tx} - P_{Tx}^{\min}}{P_{Tx}^{opt} - P_{Tx}^{\min}} & P_{Tx}^{\min} \le P_{Tx} < P_{Tx}^{opt} \\ 1 & P_{Tx} = P_{Tx}^{opt} \\ \dfrac{P_{Tx}^{\max} - P_{Tx}}{P_{\max} - P_{opt}} & P_{Tx}^{opt} \le P_{Tx} \le P_{Tx}^{\max} \\ 0 & P_{Tx} > P_{Tx}^{\max} \end{Bmatrix} \tag{22}$$

Here P_{Tx} is the calculated minimum required transmission power by a node v to reach the particular transmitter u.

4.2 Resulting degree of edge $\mu_{\tilde{G}}(E)$

In the proposed method, fuzzy relationship *min* T-norm is used to calculate the final resulting edge degree i.e. $\mu_{\tilde{G}}(E)$ of directed fuzzy graph. Fuzzy relationship Triangular

norms (T-norms) are monotone, commutative and associative functions $[0, 1] \times [0, 1] \rightarrow [0, 1]$ which are eligible for fuzzy set intersection [36].

Min T-norm has been used as during preliminary rounds of communication each node has sufficient and equal energy and the edge distance will be the sole determining factor for connection strength/weight of edge between nodes. However, in later rounds, the nodes will have non-uniform energy levels due to unequal energy consumption in each round and the residual energy shall also be another key determining factor so as to ensure to balanced energy consumption between nodes. Thus, the proposed methodology takes into consideration the dynamic nature of the network. Mathematically, for the given fuzzy graph $\tilde{G}$, the final resulting degree of edge i.e. $\mu_{\tilde{G}}(E)$ is represented as:

$$\mu_{\tilde{G}}(E_{uv}) = \underset{uv \in E}{\forall} \min(\mu(v), \mu(e_{uv})) \tag{23}$$

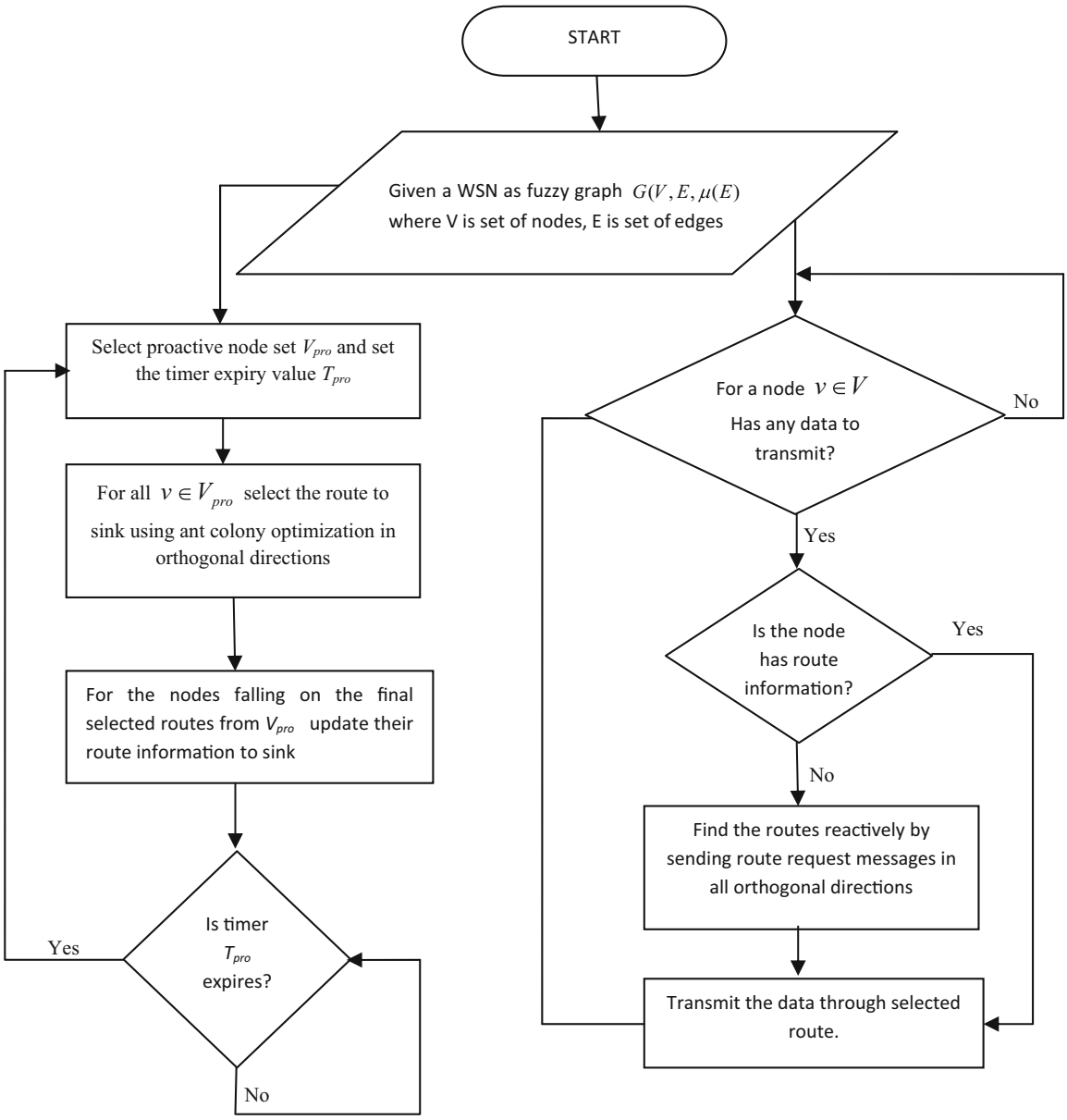


Fig. 3 Flow chart showing overview of OD-PRRP

4.3 Proposed Method "OD-PRRP"

Orthogonal Directional Proactive Reactive Routing Protocol (OD-PRRP), optimally utilizes the characteristics of both the proactive and reactive routing model. Figure 3 shows the Flow chart representing overview of the proposed method and the details associated with each element of OD-PRRP are discussed in the following sub-sections.

4.3.1 Selection of Proactive nodes

The selection of proactive nodes adopts a random back-off scheme to select certain nodes to initiate route request towards sink node. Each sensor node v sets its random back-off time b_v using a random number generator, random (0,1). At the expiry of timer, nodes select themselves as proactive nodes and send route request message in all orthogonal directions by using directional antennas. The neighboring nodes receiving this route request immediately stop their timer and thus opt out from being selected as proactive nodes. The nodes having both higher value of back-off time and which do not receive any route request till the expiry of their respective timers select themselves as proactive nodes. Thus, if any sensor node v does not receive any route request messages during its back-off time, at the expiry of its timer it considers itself as a proactive node and sends route request message in all orthogonal directions. This in turn ensures that all the nodes have accessibility to a proactive node. In other words, there is no portion of network which is left without any proactive node.

This random selection of proactive nodes using this random back-off scheme does not require any information from neighboring nodes and exchange of control messages and thus it is a simple and cost effective scheme which does not require high computations at the respective nodes. The selection of proactive nodes using back-off scheme works irrespective of size of network, deployment type and number of sensor nodes. The pseudo code for selection of *proactive* nodes is illustrated in Algorithm 1.

Algorithm.1: The pseudo-code for Proactive Nodes Selection

Procedure *PR_Selection* (V_{pro})

Initialization

1. Each ordinary Sensor node v sets its back-off time b_v using a random number generator, random(0,1),
   Where b_v is a positive real number between 0 and 1. // random back-off timer

Proactive node selection

2. ***If*** $b_v == 0$ Then
3. ***If*** sensor v does not receive any ant then //ant is route finding agent in Ant Colony Optimization
4. Select itself as proactive node
5. **end *if***
6. **end *if***
7. **end** Procedure

4.3.2 Proactive Route Discovery

The main goal of a routing protocol in wireless sensor networks is to establish an efficient route between a source and the sink. The route should be discovered and maintained with a

minimum of overhead and bandwidth consumption. In OD-PRRP, selection and utilization of some proactive nodes resolves both the issues pertaining to increased delays in reactive routing protocols and congestion losses in proactive routing protocols. The selected proactive nodes send forward ants in all the orthogonal directions. The communication in orthogonal direction resolves the necessity of global position information [5]. However, dense deployment of sensor nodes may lead to having multiple nodes in a particular orthogonal direction. In such situation, each ant chooses its next node through probabilistic method as per Eq. (24) para 4.2.3.a. Though the proactive nodes send the ants in all the orthogonal directions, each neighbor node after receiving the ant releases the ant at an angle of 180° i.e. ant continues to move in its original linear direction. In OD-PRRP each node maintains source and sender's ID along with hop count as against complete route information and is thus more scalable due to lower memory and updation requirements of state information. Also, forwarding of ants in orthogonal directions provides enhanced medium reuse. Once ants reach the sink node no more forwarding is required and depending upon the value of hop count sink updates the phenomenon value $\tau(u,v)$ using Eq. (25) and sends backward ant through the same route. This final route information shall be available with all the intermediate nodes en-route the sink reply path as well as with respective proactive nodes.

4.3.3 Ant colony optimization

Ant Colony Optimization (ACO) [8] is a meta-heuristic which works on cooperative behavior of ants with each other and is popular among researchers for solving combinational optimization problems. The basic idea behind ACO is to model the given problem as finding of best path in a graph which is representing different states of problem. These artificial ants communicate with each other by laying pheromone trails on the edges of the graph. The amount of pheromone deposition is dependent upon the quality of the solution found. The selection of the path by any ant is with respect to the probabilistic value which in turn is dependent upon the pheromone value layered by earlier ants.

In the proposed method, artificial ants find solution by moving on the given directed fuzzy graph $\tilde{G}(V, E, \mu(E)$ representing the wireless sensor network. For finding path, at each iteration the number of artificial forward ants n_a, started from a proactive node u and moving through neighbor nodes v, until ants reach at the final destination i.e. sink node. The identification of every visited node is saved onto a memory M_k, carried by ant. At each node u, selection of the next relay node v by a forward ant is based upon probabilistic decision rule which is given by:

$$P_k(u,v) = \begin{cases} \dfrac{\left[[\tau(u,v)]^{\alpha}[\lambda(u,v)]^{\beta}\right]}{\sum_{v \in N_v}\left[[\tau(u,v)]^{\alpha}[\lambda(u,v)]^{\beta}\right]} & \text{if } v \notin M_k \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

Here, $P_k(u, v)$the probability with which an ant k chooses to move from node u to v. Here $\tau(u,v)$ represents the pheromone value which depends upon the length of the path. When a forward ant reaches the destination (sink node), it is transformed into backward ant. The responsibility of backward ants is to update the pheromone trail of the path according to the desired quality of the path. Ants use reinforcement learning to discover the best paths. In reinforcement learning, the intelligent system is just given a goal to reach. The system then adopts the goal by a trial and error interaction with the environment. For

the interaction that takes the system close to the target, a positive reward is received and while going away from the target, a negative reward is assigned. All this is done by an artificial system by introducing the concept called pheromone update. The pheromone update policy depicted as follows in Eqs. (25) and (26).

$$\tau_t(u,v) \leftarrow \Delta\tau(u,v) + \rho * \tau_{t-1}(u,v) \tag{25}$$

$$\Delta\tau_s(u,v) = \begin{cases} \dfrac{1}{L_s} & \textit{if route_found} \\ 0 & \textit{otherwise} \end{cases} \tag{26}$$

where, L_s is the length of tour, which is covered by ant s at any iteration t. ρ is the pheromone decay parameter (lies between 0 to 1) which shows the evaporation rate since the last time $\tau(u, v)$ was updated. This decay parameter ensures the balanced use of resources.

$\lambda(u,v)$ is the attractiveness of the move, which can be computed by some heuristic indicating the a priori desirability of the move [40] i.e. $\lambda(u,v)$ value of the heuristic related to the performance metric under consideration. In our proposed method $\lambda(u,v)$ is made equal to fuzzy membership value $\mu_{\tilde{G}}(E)$ assigned to edges which addresses both residual energy and communication distance of the next hop node and communication is done in orthogonal directions only. α and β are the control parameters related to control the weight/ importance of pheromone and heuristic value. The pseudo code for the proposed ant colony method is given by Algorithm 2.

Algorithm.2 : ANT_ORTHO

Input: A set consisting of sensor nodes V

{Initialization phase}
Initialize
$\tau_t(u,v),\ \lambda(u,v) = \mu_G(E_{uv}),\quad \rho \quad \forall\ (u,v) \in E$
{Construction phase}

1. ∀ source u
2. best-path = ∞
3. while (n_a = no. of ants)
4. set forward ant (start in all orthogonal directions);
5. while (sink not reached y)
6. apply probability rule ( );
7. move to choose vertex in orthogonally opposite direction , Append the chosen move to the ant k's memory list M_k
8. end while
9. if length (ants journey)< length (best path)
10. best path = ants journey ;
11. end if
12. Generate backward ant
13. update pheromone;
14. End while
15. end
16. return best possible path;

4.3.4 Reactive element

In the wireless sensor networks, an event sensed by a sensor node has to be routed towards sink for further processing of the information. In case the source node is not aware of route to sink, it finds the same by sending out route request message in all locally chosen orthogonal directions. In case the neighboring node receiving the route request message also does not know the route towards sink node, it forwards it again while maintaining the similar direction i.e. to neighboring nodes at 180°. Since the source node sends the route request message (RQM) in all locally chosen orthogonal directions, it has high probability to encounter a node that has a path to the sink [5]. When the route request reaches a node with knowledge about the route to sink, no more forwarding is done and it itself sends route reply on behalf of the sink node. When the source receives the route reply message (RRM), it generates a "destination-next hop" routing entry and forwards the packets accordingly.

In case of any link failure between sink and source; the nodes reactively found other energy efficient path to the sink as discussed above. Though nodes are assumed to be densely deployed in the given region but in case if there is no neighbor in the sector covered by the required antenna, the protocol searches for neighbor in the adjacent sectors and if found it forwards the packet to such a neighbor. The allowed deviation is $\pm 90^{\circ}$ from the desired direction. The proposed method uses multiplier angle method (MAM) for deviation correction [5]. The pseudo code for the proposed method OD-PRRP is given by Algorithm 3.

Algorithm.3: OD-PRRP

Input: A set consisting of sensor nodes *SNSet*

1. Schedule each node wait time with T_{pro} delay counter
   // time to start proactive node selection after every T_{pro} time delay
2. if ($T_{pro} = 0$)
3. call Procedure PR_Selection (V_{pro}) //select proactive nodes
4. end if
5. for $\forall\ v \in V_{pro}$
6. call Procedure ANT_ORTHO // route selection using Ant Colony Optimization in orthogonal directions
7. update route information of nodes falling on
   selected path from V_{pro} to sink node v_s
8. *end for*

9. For each $v \in V$
   If *data_trx(v)* ==1 //node has data to transmit
10. if v has route information
11. Transmit data using optimal path.
12. Else
    Send *RQM* in all orthogonal directions // route request
13. Neighbor on receiving route request
14. If neighbor know the path
    Send *RRM* through the way route request received //route reply
15. Else forward the *RQM* through orthogonally opposite interface
16. End if
17. Transmit data using the discovered path
18. End if
19. End for

5 Performance Evaluation

In this section, we have evaluated the performance of proposed routing method through MATLAB simulations. For the simulation analysis, wireless sensor network is modeled into fuzzy graph. The simulation metrics and implementation are described in Sect. 5.1. On the basis of metrics discussed in 5.1, performance of the proposed method is compared with EARQ, EAODV and EEABR in Sect. 5.2. In Sect. 5.3, the performance of the proposed method is evaluated in dynamic environment by introducing node mobility. The dynamic performance has been compared with EARQ, EAODV and EEABR in terms of network lifetime, end to end delay, routing overhead and packet delivery ratio.

5.1 Simulation Environment and Metrics

We consider a wireless sensor network consisting of sensor nodes and one base station. The sensor nodes are deployed randomly in a 1000×1000 m^2 sensing field and base station is located at the center. All the nodes have an equal level of initial energy σ i.e. 2 J. Nodes in the network randomly generate multiple messages per communication period. The control and data packet sizes are fixed and kept equal to 64 bytes. To simulate, we used the log distance path loss energy model as presented in [33]. In the simulations, maximum and minimum available powers i.e. P_{max} and P_{min} are defined in terms of communication range of 200 and 80 m respectively.

For each simulated scenario, we generated ten different random topologies and for each topology we performed a simulation run consisting of 800 communication periods. For all final simulation results, each simulation is repeated ten times and a 95 % CI is obtained. Proactive node selection procedure is repeated after every ten communication periods. The anti-pheromone rate ρ is kept fixed at 0.1. In addition to this, the following system parameters as summarized in Table 1 are considered during simulations.

Performance of OD-PRRP is done in terms of: (1) Network life time: Network lifetime is defined as the distribution of alive sensor nodes with respect to number of rounds. For calculating network lifetime, we have considered the number of rounds until the first node, 25, 50 and 75 % nodes die. For each case the number of sensor nodes is varied from 100 to 500. (2) End to end transmission delay is the average of both route discovery latency and data transmission delay experienced by the source nodes while transmitting data packets. (3) Overhead in terms number of packets, which is the ratio of total number of transmitted control packets to the transmitted actual data packets. (4) Packets delivery ratio, which is the ratio of total number of packets sent to the total number of packets received. (5) Load distribution, which presents the average variance and standard deviation in resulted residual energy of nodes after every 100 communication rounds.

The simulation experimental analysis also includes the comparison of the OD-PRRP protocol against routing protocols EARQ, EAODV and EEABR. EARQ, EAODV and

Table 1 Simulation parameters

E_{elec}	100 nJ/bit
E_{amp}	50 pJ/bit/m^4
M_k	4
N	[100–500]
Route timeout	5 s
Minimum required energy(E_{min})	50 nJ

EEABR address the proactive, reactive and ant colony optimization based category of routing methods respectively. Also, compared protocols represent the energy aware approach of routing in wireless sensor networks. In EARQ, based upon the delay, reliability and energy consumption from a node to sink, probabilities are assigned to nodes proactively. Whenever actual data is to be communicated, a node selects the next forwarding neighbouring node based upon the proactively maintained probability values. EAODV is an extension of one of the basic reactive routing protocol AODV and comes out to be more competant than AODV for wireless sensor networks due to its ability to consider suitable transmission power required for transmission and total power consumption while selecting a route. EEABR is a ACO based method, which proactively forms routes from source to sink nodes. In EEABR, pheromone updation is based on the length of the route in terms of hops, average residual energy and minimum residual energy of the route. For the simulation of EARQ, EAODV and EEABR protocols, we use the same values as suggested in [9–11] for its parameters. For comparison same random seed has been used to generate the identical WSN topology.

5.2 Results

In this section, we analyze the simulation results. The comparison of the results is in terms of five parameters i.e. network life time, route discovery latency, overhead, packets delivery ratio and load distribution is discussed as under:

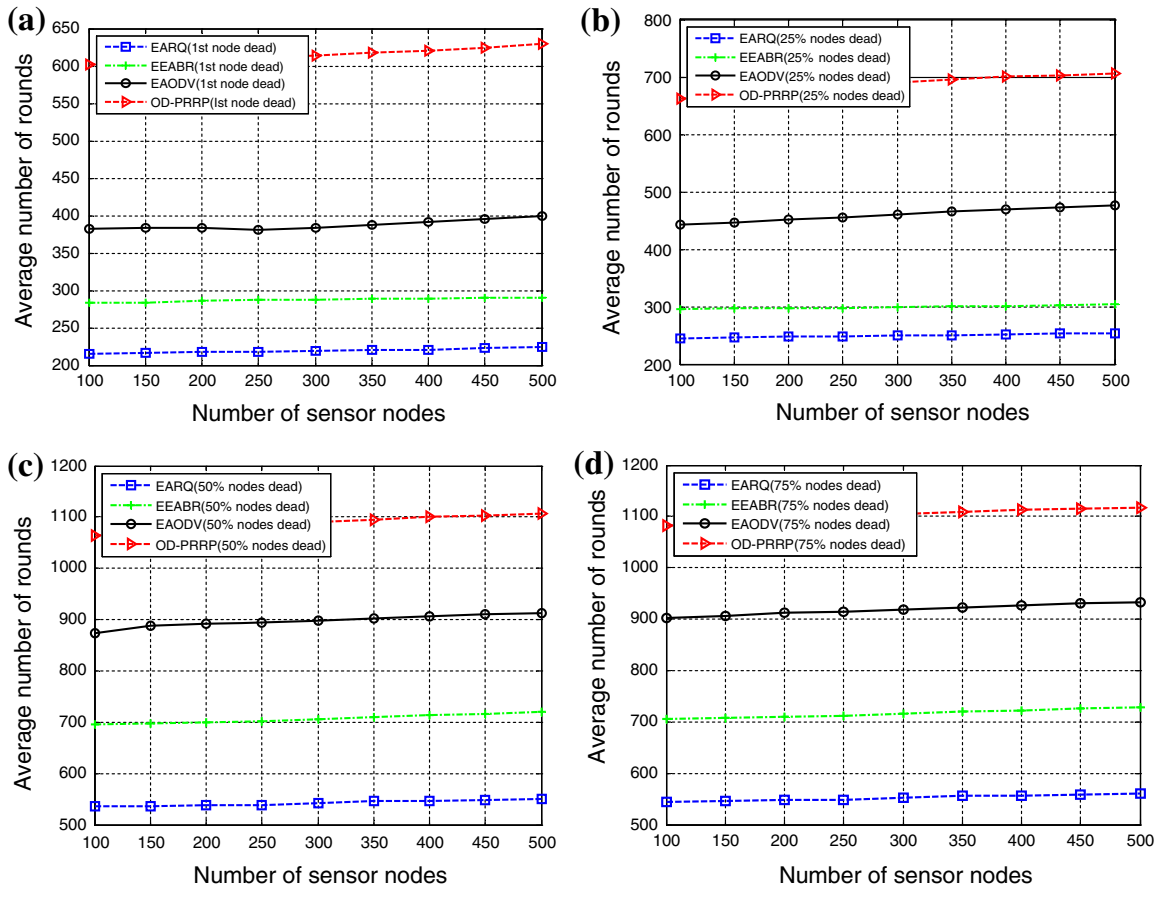


Fig. 4 Lifetime comparison between EARQ, EAODV, EEABR and proposed OD-PRRP with varying number of sensor nodes

5.2.1 Network lifetime

This metric examines the energy efficiency by calculating the number of alive nodes. The simulation results for network lifetime are shown in Fig. 4a–d. Figure 4a shows the number of communication rounds when 1st node drain off the energy. Figure 4b–d shows the number of rounds before 25, 50 and 75 % nodes of the network are dead. The network is considered operationally dead when 50 % of the nodes drain off their energy and when 75 % of the nodes drain off their energy the network is considered fully dead. For observing network lifetime, the number of sensor nodes is varied from 100 to 500. The results indicate that the better performance of our proposed method results in deferment of death of the nodes indicating thereby highest operational network life. The reasons for better performance can be attributed to the following factors:

1. Performance of EARQ and EEABR in terms of network lifetime is lowest, because it uses a pure proactive method and energy dissipated during proactive route discovery is quite high.
2. OD-PRRP has outperformed with respect to EARQ and EAODV as OD-PRRP has used orthogonal direction transmission, has characteristics of both reactive and proactive routing protocol.
3. OD-PRRP has considered expected residual energy, hop distance and required transmission power as decision parameter to select next suitable forwarder node. The ability of OD-PRRP to maintain even load distribution results in more number of functional rounds and results in improved network lifetime.

5.2.2 End to End Transmission Delay

This metric examines the average of both route discovery latency and data transmission delay experienced by the source nodes while transmitting data packets. The average end-to-end delay is measured in seconds. The route discovery latency is the time lag between the transmission of route request and receipt of route reply. Transmission delay of a data

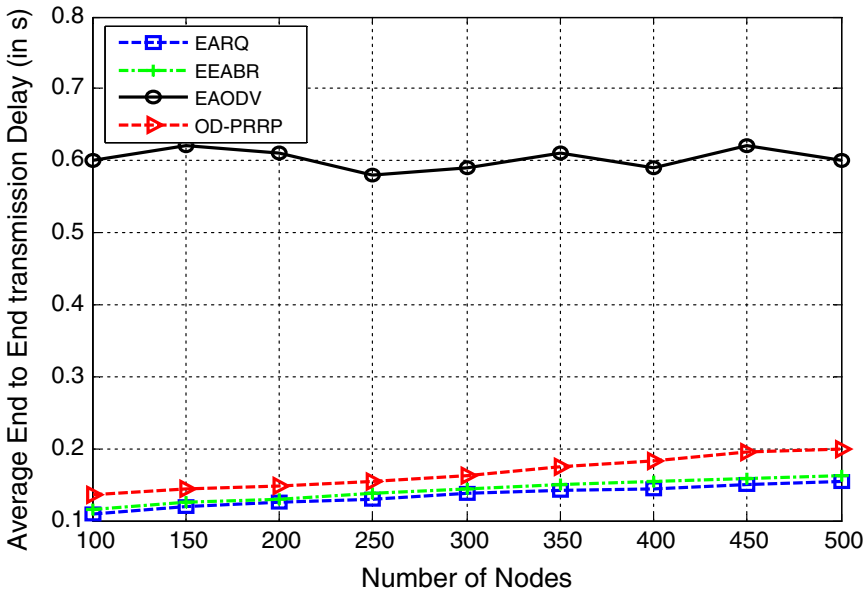


Fig. 5 Comparison of end to end delay in terms of seconds with varying number of sensor nodes

packet is the time period between transmission of data packet by source node and receipt of same at sink node.

The results for average end-to-end delay are shown in Fig. 5. From the results it has been observed that average end to end delay for EARQ, EEABR and OD-PRRP increases with increase in number of nodes. Moreover, results indicate that EARQ performs best in terms of low packet delivery delay. The resulted end to end delay in case of OD-PRRP is slightly higher than EARQ and EEABR, but the same is lower than EAODV.

5.2.3 Overhead in terms number of packets

Figure 6 shows the overhead i.e. control packets transmitted by nodes for finding routes for OD-PRRP, EARQ, EAODV and EEABR. This observation is obtained at different network densities. The Fig. 6 shows the overhead increases with increase in network density. However, for all the node densities resulted overhead in OD-PRRP is lower than EARQ, EAODV and EEABR. The observed overhead is lower due to route finding both proactively and reactively in orthogonal directions.

5.2.4 Packets delivery ratio

In Fig. 7 the performance of OD-PRRP, EARQ, EAODV and EEABR is observed in terms of packet delivery ratio by varying the number of nodes and number of flows. As shown in Fig. 7, as the number of sensor node increases, there is an increase in packet delivery ratio in the network. This increase in delivery ratio is due to increased connectivity of nodes. For simulations of packet delivery ratio, Data generation is done on the basis of probability *P*, *only if P* is greater than a certain value then node transmits a data packet to base station. Thus it is not necessary that in each round every node will generate data and there may be some nodes which are not generating data. It has been observed that packet delivery ratio in OD-PRRP is better as compared to EAODV and is comparable to EARQ and EEABR. Also number of packets successfully transmitted has improved. Improvement in delivery

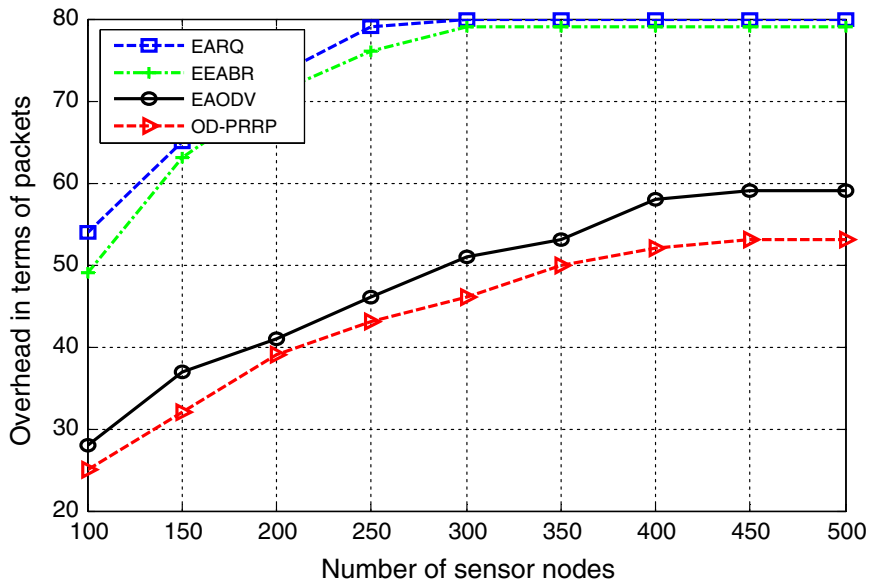


Fig. 6 Comparison of overhead in terms of packets with varying number of sensor nodes

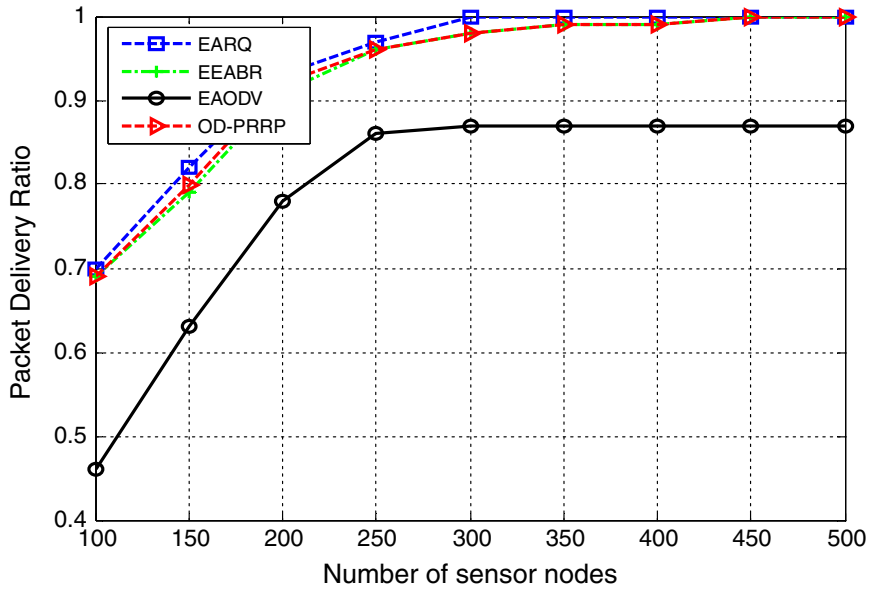


Fig. 7 Comparison of packet delivery ratio with varying Number of sensor nodes

ratio as well successful transmission of packets can be attributed to directional communication.

5.2.5 Load distribution

Load distribution calculations are done on the basis of average and standard deviation in residual energy of nodes after every 100 communication rounds. Figure 8 shows the average remaining energy of the nodes with respect to communication rounds. From the Fig. 8 is has been observed that as the number of rounds increase, OD-PRRP obtains high average remaining energy after communicating as the number of rounds increases. This is because of directional communication and low overhead. Further, it has been observed that energy consumption in OD-PRRP is more linear than other approaches, which is due to

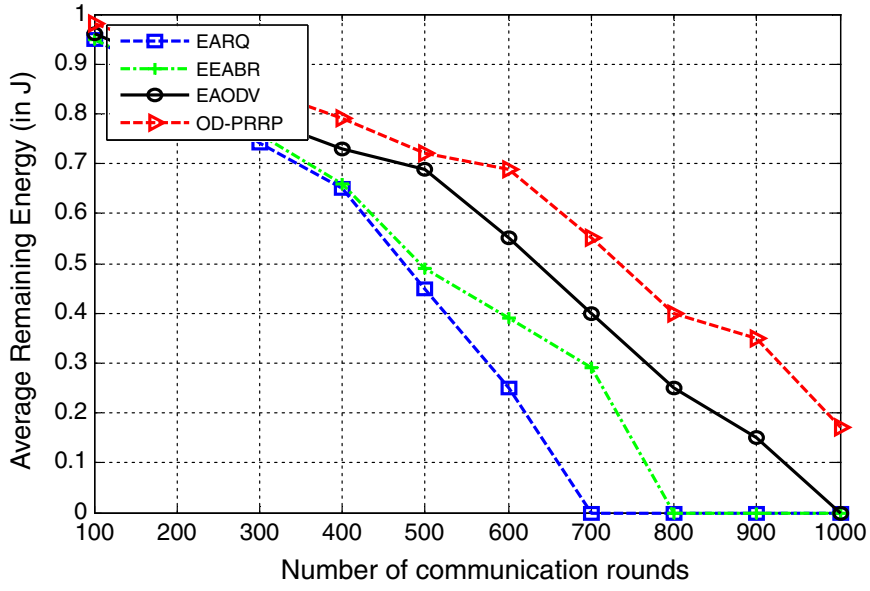


Fig. 8 Comparison of average remaining energy after every 100 communication rounds for 200 deployed sensor nodes

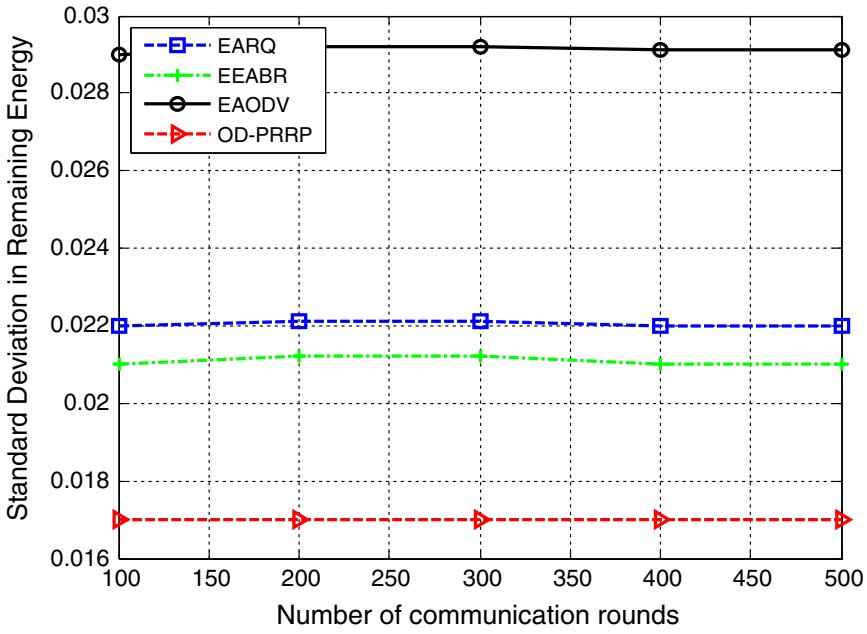


Fig. 9 Comparison of standard deviation of remaining energy after every 100 communication rounds for 200 deployed sensor nodes

orthogonal directional communication. Hence it can be inferred that energy utilization is more balanced in OD-ODRP.

In Fig. 9 we have calculated the standard deviation of the remaining energy at the end of every 100 communication rounds. Our proposed method has resulted in steady energy consumption by the sensor nodes. In OD-PRRP directional orthogonal communication with both reactive and proactive protocols, the energy consumption by sensor nodes is more uniform. This in turn makes more uniform energy consumption by sensor nodes during simulation time and in turn results in least standard deviation.

Thus the overall simulation results indicate that OD-PRRP performs better than other state of art methods i.e. EARQ, EAODV and EEABR. The same can be attributed to the orthogonal directional transmission and harnessing the benefits of both proactive and reactive methodologies.

5.3 Performance of the OD-PRRP in Dynamic Environment

In this section, we have analyzed the performance of the OD-PRRP, EARQ, EEABR and EAODV in dynamic environment i.e. after incorporating node mobility. For mobile environment, the performance has been evaluated in terms of its effect on network lifetime, packet overhead, end-to-end transmission delay and packet delivery ratio. In order to evaluate the performance of the methods for dynamic WSNs, we have used random waypoint mobility model which gives a true random node movement as expected in many mobile WSN applications. The Random Waypoint Mobility Model (RWMM) includes pauses between changes in direction and/or speed [41]. A mobile node begins by staying in one location for a certain period of time (i.e. pause). Once this time expires, the mobile node chooses a random destination in the simulation area and with a speed that is uniformly distributed between [min-speed, max-speed]. The mobile node then travels toward the newly chosen destination at this speed. Upon reaching the destination, mobile node pauses for a specified period of time before it again repeats this process.

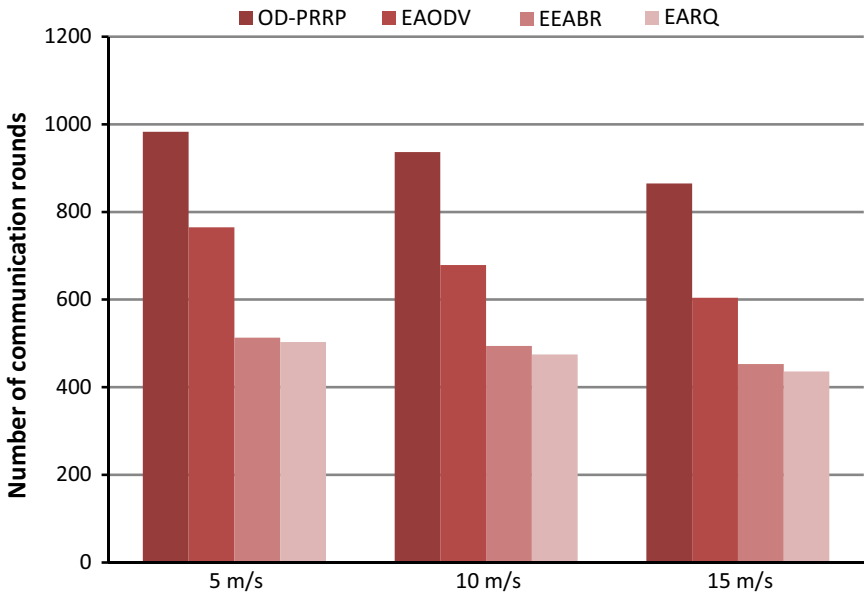


Fig. 10 Lifetime comparison between EARQ, EAODV, EEABR and proposed OD-PRRP at different node speeds

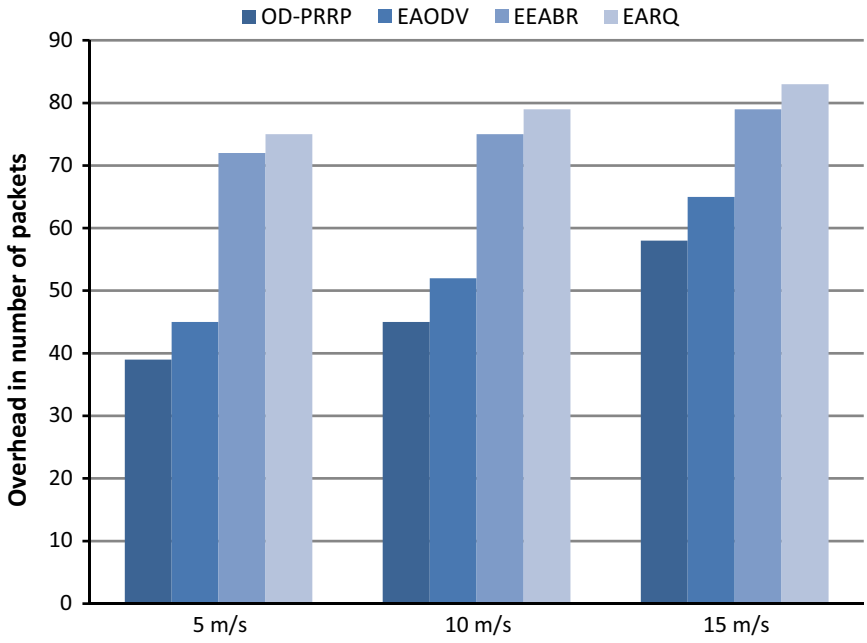


Fig. 11 Comparison of overhead in terms of packets at different node speeds, node density 200

For the simulations, route timeout parameter has been introduced and its value is set as 5 s [5, 9]. The simulation is done for a time period of 200 s, while keeping the other simulation parameters as considered in previous section. The performance of the protocols is evaluated at node density 200 and speeds viz. 5, 10 and 15 m/s are considered for node movement. These speeds have been chosen on the basis of some practical mobile WSN applications such as water quality monitoring in closed environment (low speed), health care monitoring (medium speed) or border security surveillance monitoring (higher speeds). The results calculated are based on the average of 50 trials of simulation.

Simulation results in Fig. 10 show the network lifetime of all the considered methods at different node speeds. From the analysis it has been observed that the network lifetime of

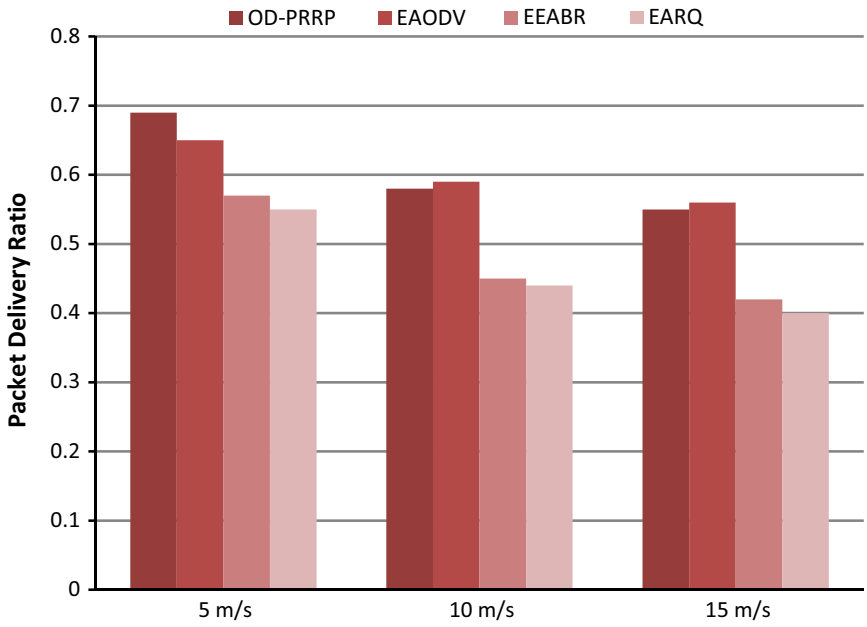


Fig. 12 Comparison of packet delivery ratio at different node speeds, node density 200

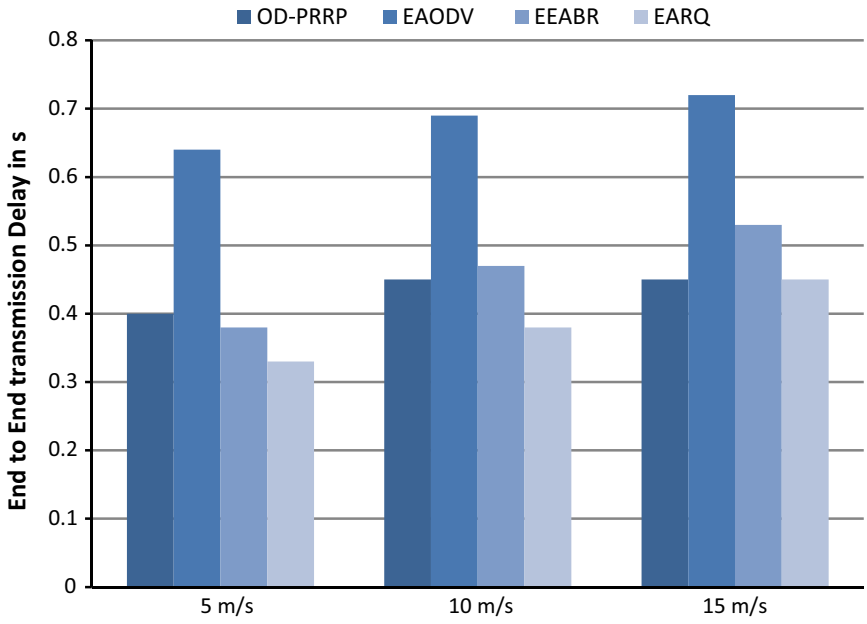


Fig. 13 Comparison of end to end delay in terms of seconds at different node speeds, node density 200

the all the methods decreased with the increase in node speed. However, the performance of OD-PRRP is better than other methods.

In respect to simulation results pertaining to overhead as shown in Fig. 11, it has been observed that in case of all the considered methods there is increase in overhead with increase in node speed. However at all the node speeds, our proposed method results low overhead as compared to other methods.

Figure 12 shows the comparison of packet delivery ratio at different node speeds. From the results it has been observed at low speeds, the packet delivery ratio of OD-PRRP,

EAODV is comparable. However at higher speeds, EARQ performs marginally better than OD-PRRP. The packet delivery ratio of EEABR and EARQ is lower at all the speeds.

Figure 13 shows that the end-to-end delay of OD-PRRP is lower than EAODV at all node speeds and is comparable to EARQ and EEABR at lower movement speeds viz. 5 and 10 m/s. From the results, it has been observed the proposed method OD-PRRP performs reasonably well in dynamic environments and can also be applied to mobile WSNs.

6 Conclusion

In this paper, we presented an ant colony optimization based proactive reactive routing protocol (OD-PRRP), which has used fuzzy graph and orthogonal directional transmission. The simulation results of our proposed protocol OD-PRRP shows reduction in control overhead without compromising at cost and energy efficiency and without requirement of global position information. In OD-PRRP, both the proactive routing and reactive routing methodology has been used to achieve fast response along with lower overhead. In the proposed method we have modeled the wireless sensor network as fuzzy edge and vertex graph. The proposed modeling by fuzzy graph has addressed the significant issue arising from the dynamic nature of the WSNs i.e. non uniform energy consumption while ensuring optimization of transmission power. The performance of OD-PRRP has been evaluated by comparing with existing available methods viz. EARQ, EAODV & EEABR and the results indicated that it performs better at all significant parameters viz. Network lifetime, packet delivery ratio, delay and load distribution. The method can be applied for design of several types of sensor networks that require energy efficiency, scalability, prolonged network lifetime, and low control overhead without requiring location information e.g. in secured battlefield surveillance, habitat monitoring, underwater monitoring, and underground coal mine monitoring. The performance of the purposed method has also been evaluated in dynamic environment. From the results, it has been observed the proposed method OD-PRRP performs reasonably well in both static and dynamic environment. In future the proposed work can be extended for optimized routing in wireless mesh and adhoc networks.

References

1. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks, 38*(4), 393–422.
2. Elson, J., & Römer, K. (2003). Wireless sensor networks: A new regime for time synchronization. *ACM SIGCOMM Computer Communication Review, 33*(1), 149–154.
3. Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks, 52*(12), 2292–2330.
4. Akkaya, K., & Younis, M. (2005). A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks, 3*(3), 325–349.
5. Cheng, B. N., Yuksel, M., & Kalyanaraman, S. (2009). Orthogonal rendezvous routing protocol for wireless mesh networks. *IEEE/ACM Transactions on Networking (ToN), 17*(2), 542–555.
6. Lee, S., Choe, H., Park, B., Song, Y., & Kim, C. K. (2011). LUCA: An energy-efficient unequal clustering algorithm using location information for wireless sensor networks. *Wireless Personal Communications, 56*(4), 715–731.
7. Perkins, C. E., & Royer, E. M. (1999, February). Ad hoc on-demand distance vector routing. In *Second IEEE workshop on mobile computing systems and applications, 1999. Proceedings. WMCSA'99*. IEEE, pp. 90–100.

8. Dorigo, M. (Ed.) (2006). *Ant colony optimization and swarm intelligence: 5th international workshop, ANTS 2006, Brussels, Belgium, September 4–7, 2006, Proceedings* (vol. 4150). Springer.
9. Heo, J., Hong, J., & Cho, Y. (2009). EARQ: Energy aware routing for real-time and reliable communication in wireless industrial sensor networks. *IEEE Transactions on Industrial Informatics, 5*(1), 3–11.
10. Li, W., Chen, M., & Li, M. M. (2009). An enhanced AODV route protocol applying in the wireless sensor networks. In *Fuzzy information and engineering volume 2*. Berlin Heidelberg: Springer, pp. 1591–1600.
11. Camilo, T., Carreto, C., Silva, J. S., & Boavida, F. (2006). An energy-efficient ant-based routing algorithm for wireless sensor networks. In *Ant colony optimization and swarm intelligence*. Berlin Heidelberg: Springer, pp. 49–59.
12. Perkins, C. E., & Bhagwat, P. (1994, October). Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *ACM SIGCOMM computer communication review*, vol. 24(4). ACM, pp. 234–244.
13. Clausen, T., Jacquet, P., Adjih, C., Laouiti, A., Minet, P., Muhlethaler, P., & Viennot, L. (2001). Optimized link state routing protocol (OLSR). In *IEEE INMIC'01*, pp. 151–162.
14. Ogier, R., Templin, F., & Lewis, M. (2003). *Topology dissemination based on reverse-path forwarding (TBRPF): Correctness and simulation evaluation*. Technical report, SRI International, Menlo Park, California.
15. Johnson, D. B., & Maltz, D. A. (1996). Dynamic source routing in ad hoc wireless networks. In *Mobile computing*. US: Springer, pp. 153–181.
16. Lee, S.-J., Gerla, M., & Chiang, C.-C. (1999, September). On-demand multicast routing protocol. In *Proceedings of IEEE WCNC*.
17. Yu, J. Y., Chong, P. H. J., & Zhang, M. (2008, September). Performance of efficient CBRP in mobile ad hoc networks (MANETS). In *Vehicular technology conference, 2008. VTC 2008-Fall. IEEE 68th*. IEEE, pp. 1–7.
18. Karp, B., & Kung, H. T. (2000, August). GPSR: Greedy perimeter stateless routing for wireless networks. In *Proceedings of the 6th annual international conference on mobile computing and networking*. ACM, pp. 243–254.
19. Beijar, N. (2002). *Zone routing protocol (ZRP)*. Finland: Networking Laboratory, Helsinki University of Technology.
20. Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., & Silva, F. (2003). Directed diffusion for wireless sensor networking. *IEEE/ACM Transactions on Networking, 11*(1), 2–16.
21. Sohrabi, K., Gao, J., Ailawadhi, V., & Pottie, G. J. (2000). Protocols for self-organization of a wireless sensor network. *IEEE Personal Communications, 7*(5), 16–27.
22. Chang, J. H., & Tassiulas, L. (2004). Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Transactions on Networking (TON), 12*(4), 609–619.
23. Li, G., Wang, L., & Li, H. (2008, April). Multiple-objective fuzzy decision making based routing protocol for wireless sensor networks. In *IEEE international conference on networking, sensing and control, 2008. ICNSC 2008*. IEEE, pp. 1273–1278.
24. Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing, 3*(4), 366–379.
25. Minhas, M. R., Gopalakrishnan, S., & Leung, V. C. (2009, April). An online multipath routing algorithm for maximizing lifetime in wireless sensor networks. In *Sixth international conference on information technology: New generations, 2009. ITNG'09*. IEEE, pp. 581–586.
26. Jabbar, S., Minhas, A. A., Akhtar, R. A., & Aziz, M. Z. (2009, December). REAR: Real-time energy aware routing for wireless adhoc micro sensors network. In *Eighth IEEE international conference on dependable, autonomic and secure computing, 2009 (DASC'09)*. IEEE, pp. 825–830.
27. AlShawi, I. S., Yan, L., Pan, W., & Luo, B. (2012). Lifetime enhancement in wireless sensor networks using fuzzy approach and A-star algorithm. *IEEE Sensors Journal, 12*(10), 3010–3018.
28. Ortiz, A. M., Olivares, T., Castillo, J. C., Orozco-Barbosa, L., Marrón, P. J., & Royo, F. (2010, October). Intelligent role-based routing for dense wireless sensor networks. In *2010 third joint IFIP wireless and mobile networking conference (WMNC)*. IEEE, pp. 1–6.
29. Ahvar, E., Pourmoslemi, A., & Piran, M. J. (2011). Fear: A fuzzy-based energy-aware routing protocol for wireless sensor networks. arXiv preprint arXiv:1108.2777.
30. Ding, N., & Liu, P. X. (2005). A centralized approach to energy-efficient protocols for wireless sensor networks. In *2005 IEEE international conference on mechatronics and automation* vol. 3. IEEE, pp. 1636–1641.
31. Wen, Y. F., Chen, Y. Q., & Pan, M. (2008). Adaptive ant-based routing in wireless sensor networks using Energy* Delay metrics. *Journal of Zhejiang University SCIENCE A, 9*(4), 531–538.

32. Liu, Y., Zhu, H., Xu, K., & Jia, Y. (2007, August). A routing strategy based on ant algorithm for WSN. In *Third international conference on natural computation, 2007. ICNC 2007, v*ol. 5. IEEE, pp. 685–689.
33. Rappaport, T. S. (1996). *Wireless communications: Principles and practice* (Vol. 2). New Jersey: Prentice Hall PTR.
34. Koczy, L. T. (1992). Fuzzy graphs in the evaluation and optimization of networks. *Fuzzy sets and systems, 46*, 307–319.
35. Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338–353.
36. Ross, T. (2004). *Fuzzy logic with engineering applications* (2nd ed.). Chichester: Wiley.
37. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An Application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications, 1*(4), 660–670.
38. Kheireddine, M., & Abdellatif, R. (2014). Analysis of hops length in wireless sensor networks. *Wireless Sensor Network, 6*(6), 109–117.
39. Correia, L. H., Macedo, D. F., Silva, D. A., dos Santos, A. L., Loureiro, A. A., & Nogueira, J. M. S. (2005, October). Transmission power control in MAC protocols for wireless sensor networks. In *Proceedings of the 8th ACM international symposium on modeling, analysis and simulation of wireless and mobile systems*. ACM, pp. 282–289.
40. Maniezzo, V., & Carbonaro, A. (2002). Ant colony optimization: An overview. In *Essays and surveys in meta heuristics*. Springer US, pp. 469–492.
41. Camp, T., Boleng, J., & Davies, V. (2002). A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing, 2*(5), 483–502.



Aarti Jain graduated with a B.Tech degree in Electronics and Communication Engineering from Beant Govt. Engineering College, Punjab, India in June 2002. Then she obtained her M.E. degee in Electronics and Communication from Delhi College of Engineering, University of Delhi, India in 2009. Currently, she is working as Asst. Prof. at the Department of Electronics and Communication Engineering, Ambedkar Institute of Advanced Communication Technology and Research. She is currently working towards the PhD degree from Guru Gobind Singh Indraprastha University, Delhi. Her currentresearch interests are in Energy and quality management for wireless sensor networks.



B.V. Ramana Reddy is Professor of USICT G.G.S. Indraprastha University, Dwarka. He obtained his ME and Ph. D (ECE) degrees during 1991 and 1998 respectively. He served in various capacities and has about 20 years of experience in teaching and research. Before joining this University he served as Asstt. Professor at NIT Hamirpur. He has about 25 publications in journals and International Conferences to his credit. He is a Fellow of IETE and a member of other professional bodies such as IE, IEEE, ISTE and SEMCEI. He has guided few scholars leading to M. Tech and Ph.D's. He has currently 6 students registered for Ph. D. His current research interest includes Wireless communications which include mobile, Adhoc and sensor based networks, computer communication networks, Semiconductor and VLSI circuits and microwave & optical communications

Global Journal of Flexible Systems Management (June 2015) 16(2):115–132
DOI 10.1007/s40171-014-0085-6

ORIGINAL ARTICLE

Assessment of Critical Enablers for Flexible Supply Chain Performance Measurement System Using Fuzzy DEMATEL Approach

Mohit Tyagi · Pradeep Kumar · Dinesh Kumar

Received: 20 June 2014 / Accepted: 28 October 2014 / Published online: 26 November 2014
© Global Institute of Flexible Systems Management 2014

Abstract *In the competitive world of today, each organization has a desire to sustain in the marketplace with the implementation of their healthier and flexible supply chain performance measurement (SCPM) system. For the successful implementation, needs to know the significant set of enablers. This study identifies a set of important enablers based on literature review and discussion with field experts of automobile manufacturing industries located in the National Capital Region of India. The vagueness and impreciseness of field expert's judgements has been reduced using fuzzy decision making trial and evaluation laboratory (fuzzy DEMATEL) approach and analyzed the enablers in order to implement a flexible SCPM system. The findings of this research advocate that enabler 'higher customer satisfaction' comes in picture with highest value of 'Prominence' (6.4272) and 'Relation' (1.0354), therefore seems as a most significant and influencing enabler, while on other side the enabler 'proper capacity utilization' is considered as ample influencing enabler, because it has lowest Prominence' (4.4735) and 'Relation' (minus 0.9680) values. This research discussed the categorization into the cause and effect group, degree of interaction and inter-relationship of considered enablers. The outcomes of this study may provide an aid to the managers to implement an effective and flexible SCPM system through which overall profitability of an organization may be improved.*

Keywords Enablers · Fuzzy DEMATEL · Kappa statistics · Performance measurement system · Supply chain management

M. Tyagi (✉) · P. Kumar · D. Kumar
Department of Mechanical and Industrial Engineering, Indian Institute of Technology Roorkee, Roorkee 247667, Uttarakhand (UK), India
e-mail: mohitmied@gmail.com

P. Kumar
e-mail: kumarfme@iitr.ernet.in

D. Kumar
e-mail: dinesfme@iitr.ernet.in

Introduction

Since 1980s, it has been seen that supply chain management (SCM) is working as a base pillar in enhancing the overall performance of organizations. Womack and Jones (2005) found that SCM works as a key component to make the organization more effective and efficient under the competitive environment. According to Bowersox et al. (2010), aim of SCM should be to maximize the supply chain outcomes, by providing appropriate service level to the end users. Without proper understanding of SCM, it is not thinkable to construct a seamless supply chain. Therefore, there is a need to optimize the supply chain performance in order to provide the products and services at faster rate with lowest possible cost (Gunasekaran et al. 2001). That's why; to make supply chain more effective and efficient; most of the companies are inserting additional efforts towards implementation of flexible supply chain performance measurement (SCPM) system, today.

In this paper, our focus is to provide the key enablers to mangers in aspect of implementing and improving their SCPM system through incorporating proper flexibility to enhance the overall performance of an organization. In broad aspect, performance measurement can be defined as a manner of enumerating the efficacy and proficiency of activities (Neely et al. 1995). Waggoner et al. (1999) says that ample performance measurement provides a path to

identify the area that required more attention and also helpful in improving the communications level among the supply chain associates. Simatupang and Sridharan (2002) stated that a continual improvement for supply chain members, end customers and outside stakeholders could be achieved easily through an effective performance measurement system. Charan et al. (2008) says that higher effectiveness of SCPM system provides more improvement in the supply chain performance. In the previous studies many performance based frameworks were developed, for example: Qureshi et al. (2008) developed an integrated framework using interpretive structural modeling (ISM) and fuzzy MICMAC analysis to examine the key criteria of 3PLs providers in order to enhance the effectiveness of SCM. Thakkar et al. (2009) proposed an integrated SCPM framework for small and medium scale enterprises (SMEs) using a set of qualitative and quantitative insights. Mangla et al. (2014) analyzed a sustainable supply chain performance based framework. Sidola et al. (2012) analyzed the impact of information technology (IT) capabilities on an organization's performance using system dynamics concept. Tyagi et al. (2014) developed and analyzed the IT enabled supply chain performance model for Indian automobile industry located at Delhi region.

During the competitive environment, customer not only oriented towards high product quality with lower pricing but also expects customized products with short delivery time. To meet such type of expectations, flexible systems management may be helpful (Kumar and Deshmukh 2006). Wadhwa et al. (2007) developed the various flexibility based performance models related to product recovery programs and visualized the importance of flexible system in decision making process. Hu et al. (2013) examined the performance of centralized supply chain by incorporating the flexible ordering policies. Thome et al. (2014) discussed the key effects of supply chain flexibility capabilities at multi-tiers level that required in delivering the products from point of origin to the point of consumptions. Sushil (2012) stated that the flexible systems management is quite valuable in gaining and improving the performance of supply chain operations as a whole.

From last two decades, Indian automobile sector is attracting the automobile manufacturers like: Maruti Suzuki, Hyundai, Honda, Toyota, Nissan etc. to set up manufacturing bases at different locations in India. It have amplified the competition among the manufacturers and encouraging them to make a product at low cost with better quality in order to fulfill the customer's demand. To sustain in this competitive environment, most of the automobile manufacturers are trying to improve their SCPM system in order to build a seamless supply chain to fulfill the customer's requirements. Many previous studies have discussion on supply chain performance measurement in the various sectors; however, there is an absence of analysis of performance matrices or barriers/enablers (Gunasekaran et al. 2004). Saad and Patel (2006) found that there is a limited research that concentrates on the supply chain performance system of an Indian automobile manufacturing company. By keeping above themes in the mind, automobile manufacturing companies located at NCR of India have been chosen for this research.

This study aims to identify a set of enablers based on literature review and field expert's opinions, and to analyze their mutual importance using fuzzy decision making trial and evaluation laboratory (fuzzy DEMATEL), a multi-criteria decision making approach. This study discussed the cause and effect behavior and inter-relationship mapping of considered enablers and addresses the following research queries:

- What is the list of significant enablers related to flexible SCPM system?
- How to address the internal consistency of identified enablers with general perspectives of an organization?
- What are the behaviour, importance ranking and interaction level of considered enablers?
- Why are ranking and interaction level of these enablers valuable for SCPM system implementation?
- How are these enablers useful for managers to improve and implement a flexible SCPM system?

Enabler's Identification: Literature Review

For a company, customer satisfaction is very essential element to improve their supply chain performance and overall profit. To make a positive customer voice, higher level of customer satisfaction is required that in turn also increases the market share/value. Moreover, it may be expected that, increase in the market share will provide a higher growth in the market place. As an outcome of this, it is easy for the company to provide its products at a minimum price with better quality and will also improve the overall performance of a company (Anderson et al. 1994).

There are many aspects to improve the production performance; out of them proper capacity utilization can be selected as a significant measure to enhance the supply chain performance. The capacity utilization has a direct impact on lead time, flexibility and delivery system and also affects the customer demand by enhancing the speed of response (Slack et al. 1995). According to Christopher (1998), the company can achieve a competitive advantage on the basis of proper focusing towards their service and cost leadership. For this, it is required to build a flexible and appropriate performance measurement of supply chain to promote the better understanding among the supply chain stages (Fawcett and Cooper 1998; Dreyer 2000).

Today's most of the organizations are trying to enhance the use of Information Technologies (IT) in their SCM system, to make it more efficient and effective. It indicates that IT works as a supply chain enabler and helps to the managers in developing information systems and in providing more flexibility among the work related tasks. IT provides a seamless flow of entities among the existing stages of supply chain and also built the organization functions more efficient throughout the whole supply chain (Handfield and Nichols 1999). Finally, information technology can be considered as a key enabler/driver in implementing a flexible SCPM system (Koh et al. 2006). Tompkins and Ang (1999) stated that effectively and timely use of information among the supply chain stages works as a key enabler/driver for this competitive market place. According to Lin and Tseng (2006) appropriate level of information sharing among the supply chain partners provides the reduction in total cost, improvement in order fulfillment and reduction in order cycle time, thus as a result improving the overall supply chain performance.

Ragatz et al. (1997) say that effective integration among the suppliers and manufacturers plays a very important role in improving the product value or supply chain performance to remain competitive in the market place. The various existing theories suggested that if degree of integration among the supply chain stages is higher, then it is necessary that firm performance will also be better (Anderson and Katz 1998; Narasimhan and Jayaram 1998; Johnson 1999; Frohlich and Westbrook 2001). Barut et al. (2002) found that the inter-company integration and coordination can be achieved by proper information sharing and works as a significant success factor to improve the supply chain performance. Turner (1993) says that supply chain integration helps to make an effective communication and flexibility among the all supply chain stages. With this continuation, it has been seen that level of information sharing and information suitability are important to achieve the customer responsiveness. Daugherty et al. (1995) support the above statement and says that information availability and customer responsiveness are positively related to each other in order to improve the supply chain performance of an organization.

Green and Inman (2005) noticed that alignment of systems in all supply chain stages ranging from point of origin to the point of consumption will help to enhance the organizational performance. To obtain competitive advantage at the supply chain phase, marketing strategies should be appropriately implemented and executed with a higher level of integration and coordination, to concentrate more towards the end consumer (Mavi et al. 2013). To achieve the competitive advantage, supply chain performance also depends on its partner's talent under dynamic environment (Vanderhaeghe and de Treville 2003).

Gunasekaran and Kobu (2007) say that open and transparent communication between people builds a better co-operative and supportive work environment. It also facilitates to develop an appropriate set of performance measures in order to make a SCPM system more flexible and efficient, by which organizational performance will improve very rapidly. During this competitive and changing behavior of market, it is required to focus towards the new product design and development. As an advancement of manufacturing technologies and innovations throughout the prominent organizations, product development can be seen as a significant source of competitive advantage. Therefore, it is mandatory to reduce the new product development (NPD) cycle time in order to make an improvement in market share and profit, and to remain stable in this competitive environment (Cooper and Kleinschmidt 1995; Griffin 1997; Ittner and Larcker 1997). An organization can also improve its performance by adopting the modern manufacturing practices such as quality management, just-in-time and information technology (Flynn and Flynn 2005; Green and Inman 2005). Little et al. (1995) noticed that effectiveness of scheduling techniques directly affects the production system and also helps in improving the supply chain performance.

According to Heizer and Render (2006) an effective SCM should have the ability to sustain long-term strategic relationships with their supply chain associates. Because, the quality of supply chain relationships have direct impact on the supply chain performance (Fynes et al. 2004). Kohli and Jaworski (1990) found that improvement in organizational performance can also be achieved by adopting the market orientation. Hence, adoption of market orientation will play a crucial role in improving the supply chain performance.

Research Methodology

The objective of this research is to evaluate the importance level of considered enablers to implement a flexible SCPM system. To achieve the objective, a set of significant enablers which affects the flexible SCPM system of an organization, was identified based on literature review and discussion with field experts, chosen from automobile industry located at national capital region (NCR) of India. For the discussion purpose, extensive brainstorming sessions (with five industrial and two academic experts) were fixed to strengthen and improve the literature inputs.

As a pre-testing part, a questionnaire has been formulated on five point Likert scale (1 for least important…..5 for absolutely most important) and send to the 150 field experts to collect their opinions in order to know the importance level and to test the feasibility of enablers. For



Table 1 List of identified enablers

Sr. no.	Enablers	References
1.	Higher customer satisfaction	Klein and Roth (1993), Anderson et al. (1994), Fawcett and Cooper (1998), Beamon (1999), Ulaga and Chacour (2001), Gunasekaran et al. (2001), Rabelo et al. (2007) and Jakhar and Barua (2014)
2.	Proper capacity utilization	Slack et al. (1995), Aviv (2007) and Bhagwat and Sharma (2007)
3.	Making proper performance measurement	Christopher (1998), Fawcett and Cooper (1998), Dreyer (2000), Gunasekaran et al. (2004), Hofmann and Locker (2009) and Chan et al. (2014)
4.	Information sharing	Handfield and Nichols (1999), Tompkins and Ang (1999), Koh et al. (2006) and Lin and Tseng (2006)
5.	Effective integration of suppliers	Morgan and Monczka (1996), Ragatz et al. (1997), Das et al. (2006), Harrtman and Bengtsson (2009) and Lockstrom et al. (2010)
6.	Higher degree of integration across the supply chain	Turner (1993), Anderson and Katz (1998), Narasimhan and Jayaram (1998), Johnson (1999), Frohlich and Westbrook (2001), Barut et al. (2002) and Petersen et al. (2005)
7.	Information availability and customer responsiveness	Daugherty et al. (1995), Narasimhan and Jayaram (1998), Van der Vorst and Beulens (2002), Ravi et al. (2005), Holweg et al. (2005) and Lin and Chang (2008)
8.	Alignment of systems throughout the supply chain	Green and Inman (2005), Cousins (2005), Brown et al. (2007) and Mavi et al. (2013)
9.	Open and transparent communication	Gupta and Govindarajan (2000), Kaynak (2003), Gunasekaran and Kobu (2007) and Gunasekaran and Ngai (2008)
10.	Reduction of new product development cycle time	Cooper and Kleinschmidt (1995), Griffin, (1997), Ittner and Larcker (1997), Labahn et al. (1997), Mishra and Shah (2009) and Lawson and Potter (2012)
11.	Adoption of modern manufacturing practices	Fawcett and Myers (2001), Flynn and Flynn (2005), Ketokivi and Schroeder (2004) and Green and Inman (2005)
12.	Effectiveness of scheduling techniques	Little et al. (1995), Gunasekaran et al. (2004), Bhagwat and Sharma (2007), Giannakis (2011) and Cho et al. (2012)
13.	Long-term, strategic relationships	Graham et al. (1994), Heizer and Render (2006), Chiadamrong and Prasertwattana (2006), Paulraj et al. (2008) and Oke et al. (2013)
14.	Quality of the supply chain relationships	Lalonde (1998), Fynes et al. (2004), Kwon and Suh (2004) and Nyaga et al. (2010)
15.	Adoption of a market orientation	Kohli and Jaworski (1990), Hurley and Hult (1998), Dawes (1999), Lings (2004), Green et al. (2005) and Demirbag et al. (2006)

this purpose, a reliability analysis has been performed. After that the consistency (post-testing) of considered enablers has been checked using Kappa statistics. To derive this statistics, the opinion of twenty five experts (fifteen from industry and ten from academia) were collected using a questionnaire study, regarding the existence of fifteen enablers under the four perspectives (managerial, information, integration and production). Based on collected opinions, Kappa index value has been calculated, which gives the evidence about the internal consistency of enablers under these perspectives. In this case, Kappa index value comes under the significant range as interpreted in Tables 4 and 5. Hence, the identified set of enablers as shown in Table 1 is consistent and ready for further analysis.

After finding the consistency evidence, the set of enablers has been analyzed using fuzzy DEMATEL approach. To derive the fuzzy DEMATEL analysis, a questionnaire has been structured on linguistic terms (no influence… very high influence) and discussed with the nineteen field experts to collect their individual opinions regarding importance level of considered enablers. Then, step by step procedure of fuzzy DEMATEL approach has been adopted to meet the specified objective of present research.

Data Collection and Analysis

Reliability Analysis

With the purpose of pre-testing the feasibility of considered enablers, a questionnaire was structured and sent to the 150 professionals, out of them 57 responses were received. The response rate was found 38 % which is sufficient to drive analysis (Malhotra and Grover 1998). To check the reliability of collected data, Cronbach's alpha coefficient (α) has been computed using SPSS-16, which comes as 0.785 and is given in Table 2, i.e. under the recommended range ($0.7 < \alpha > 0.95$). The mean and standard deviation are also calculated as shown in Table 3. In a mode of supply chain performance improvement, this table gives an initial picture about the importance level of enablers based on

Table 2 Reliability statistics for the considered enablers

Cronbach's alpha	Cronbach's alpha based on standardized enablers	No. of enablers
0.785	0.784	15

their mean value. For example, enablers such as higher customer satisfaction, quality of the supply chain, relationships, and alignment of systems throughout the supply chain have higher mean values in comparison to the other enablers. Hence, it is an indication that these enablers have more importance in edifice of better and flexible supply chain performance system.

Check for Internal Consistency: Kappa Statistics

The kappa statistics developed by Cohen (1968) has been applied to check the consistency (post-testing part) of considered enablers under the four categories namely as: C_1 = managerial perspective; C_2 = information perspective; C_3 = integration perspective and C_4 = production perspective. As a broad aspect, the identified enablers can be assumed under the above four perspectives of an organization, thus to know the nature of consistency of each enabler under the corresponding perspective, kappa statistics has been used in this study.

The kappa index (k) is known for true and actual consensus assessment and defined as the ratio of difference of observed theoretical and chance agreement to the beyond by chance agreement. It is a quantitatively rated contract rather than predicted by chance and is given as:

$$k = \frac{proportion\,of\,the\,observed\,agreement - chance\,agreement}{1 - chance\,agreement} = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}, \tag{1}$$

where

$$\bar{P} = \frac{1}{u}\sum_{i=1}^{u} P_i \tag{2}$$

$$\bar{P}_e = \sum_{j=1}^{v} P_j^2 \tag{3}$$

The assignments percentage P_j under the jth category and degree of involvement (P_i) of raters/experts for the ith enabler are calculated by using the Eqs. 4 and 5, as given below:

$$P_j = \sum_{i=1}^{u} w_{ij} \Big/ \sum_{j=1}^{v}\left(\sum_{i=1}^{u} w_{ij}\right) \tag{4}$$

$$P_i = \frac{1}{W(W-1)}\sum_{j=1}^{v}\left(w_{ij}^2 - W\right) \tag{5}$$

The notations used in the Eqs. 2–5 are described as:

u = Total number of identified enablers = $\sum A_m$, $(m = 1, \ldots 15)$

v = Total number of categories = $\sum C_n$, (n = 1, ….4)

W = Total number of considered experts = 25

w_{ij} = Ratings of ith enabler with respect to jth category, $(i = 1, \ldots u)$, $(j = 1, \ldots v)$

Table 3 Statistics for enablers

A_m	Enablers	Mean	S. D
A_1	Higher customer satisfaction	4.2982	0.8229
A_2	Proper capacity utilization	3.6140	1.0308
A_3	Making proper performance measurement	3.5789	0.9626
A_4	Information sharing	3.4561	1.1350
A_5	Effective integration of suppliers	3.5439	1.0534
A_6	Higher degree of integration across the supply chain	3.4035	1.0996
A_7	Information availability and customer responsiveness	3.4912	1.2553
A_8	Alignment of systems throughout the supply chain	3.7895	1.0130
A_9	Open and transparent communication	3.2632	1.4333
A_{10}	Reduction of New Product Development cycle time	3.4912	1.1820
A_{11}	Adoption of modern manufacturing practices	3.7544	1.0399
A_{12}	Effectiveness of scheduling techniques	3.7368	0.9548
A_{13}	Long-term, strategic relationships	3.6316	1.0628
A_{14}	Quality of the supply chain relationships	3.9825	0.8962
A_{15}	Adoption of a market orientation	3.5965	1.0667

Table 4 Summary of 25 experts measuring 15 enablers

A_m	C_1	C_2	C_3	C_4	P_i
A_1	19	3	3	0	0.4631
A_2	17	2	5	1	0.3635
A_3	21	2	2	0	0.5793
A_4	2	19	4	0	0.4665
A_5	2	3	20	0	0.5196
A_6	2	4	19	0	0.4665
A_7	3	17	4	1	0.3569
A_8	4	3	17	1	0.3569
A_9	4	15	6	0	0.2938
A_{10}	1	1	3	20	0.5163
A_{11}	4	3	0	18	0.4133
A_{12}	18	3	3	1	0.4034
A_{13}	20	1	4	0	0.5262
A_{14}	4	3	18	0	0.4133
A_{15}	2	18	5	0	0.4200
P_j	0.3280	0.2587	0.3013	0.1120	

Table 4 has been prepared and the filled value in each cell represents the number of experts involved for the existence of enablers in corresponding category.

The value of kappa (k) is calculated as:

$$k = \frac{(0.4372 - 0.2778)}{(1 - 0.2778)} = \frac{0.1594}{0.7222} = 0.2207 \tag{6}$$

The interpretation scale proposed by Landis and Koch (1977) of kappa value is given as in Table 5.

Analyzing 'k', it is observed that it comes in the showing fair consideration of all enablers in their category. Further analysis has been done using fuzzy DEMATEL approach as given in flow chart of the work (Fig. 1).

Fuzzy Set Theory and Triangular Fuzzy Numbers

For decision making problems related to complex system, most of the organizations have formulated a questionnaire based on linguistic terms instead of crisp values and send to the field experts to gather their opinions regarding enablers. The states of each linguistic variable are represented by linguistic terms. The linguistic terms used are: very small, small, medium, large, and very large. The judgements of decision makers in terms of linguistic expressions have some degree of vagueness or fuzziness. Fuzzy set theory propose by Zadeh (1965), has been used to reduce the fuzziness associated with the human judgements, which deals with the linguistic variables to resolve the real world problems. A fuzzy set gives a degree of belongingness by incorporating the membership function between 0 and 1, while crisp set by binary logic 0 or 1 only, i.e. fuzzy set is advantageous over the crisp set to reduce the uncertainty.

Let X be a universe of discourse, $X = \{x_1, x_2, \ldots, x_n\}$, A fuzzy set $\tilde{a}$ of X is a set of order pairs like $\{(x_1, \mu_{\tilde{a}}(x_1)), (x_2, \mu_{\tilde{a}}(x_2)), \ldots, (x_n, \mu_{\tilde{a}}(x_n))\}$, where $\mu_{\tilde{a}} : X \to [0, 1]$ is the membership function of $\tilde{a}$ and $\mu_{\tilde{a}}(x_i)$ terms a membership degree of x_i in $\tilde{a}$. If $\tilde{a} \to 1$, then the grade of membership of x_i in $\tilde{a}$ will be higher (Kaufmann and Gupta 1991). Then in the universe of discourse 'X', a fuzzy set $\tilde{a}$ is denoted as:

$$\tilde{a} = \left\{ \frac{\mu_{\tilde{a}}(x_1)}{x_1} + \frac{\mu_{\tilde{a}}(x_2)}{x_2} + \frac{\mu_{\tilde{a}}(x_3)}{x_3} \ldots \right\} = \left\{ \sum_i \frac{\mu_{\tilde{a}}(x_i)}{x_i} \right\} where\, x_i \in X \tag{7}$$

(When X is discrete and finite).

$$\tilde{a} = \left\{ \int \frac{\mu_{\tilde{a}}(x)}{x} \right\} where\, x \in X \tag{8}$$

(When X is continuous and infinite).

To maximize the real world applications and ease in computation, triangular fuzzy number (TFN) has been used (Yeh and Deng 2004). Its membership function $\mu_{\tilde{a}}(x)$ has been as:

$$\mu_{\tilde{a}}(x) = \begin{cases} 0, & x < a_1 \\ \frac{x - a_1}{a_2 - a_1}, & a_1 \le x \le a_2 \\ \frac{a_3 - x}{a_3 - a_2}, & a_2 \le x \le a_3 \\ 0, & x > a_3 \end{cases}, \tag{9}$$

where in the Eq. 9, the symbols used are:

Triangular fuzzy number, $\tilde{a}$ = triplet (a_1, a_2, a_3)

a_1 = lower value of triplet, a_2 = middle value triplet and a_3 = upper value triplet.

The two TFN's (Fig. 2) $\tilde{a} = (a_1, a_2, a_3)$ and $\tilde{b} = (b_1, b_2, b_3)$ have their usual operation law and used in the analysis.

In the study, the opinions of field experts with certain degree of uncertainty have been tabulated, aggregated and searched the crisp values using CFCS (Converting Fuzzy data into Crisp Scores) defuzzification method (Opricovic and Tzeng 2003).

Table 5 Scale for strength interpretation of kappa statistics

Values of k	<0	0.1–0.20	0.21–0.40	0.41–0.60	0.61–0.80	0.81–1.00
Interpretation	Poor	Slight	Fair	Moderate	Substantial	Perfect

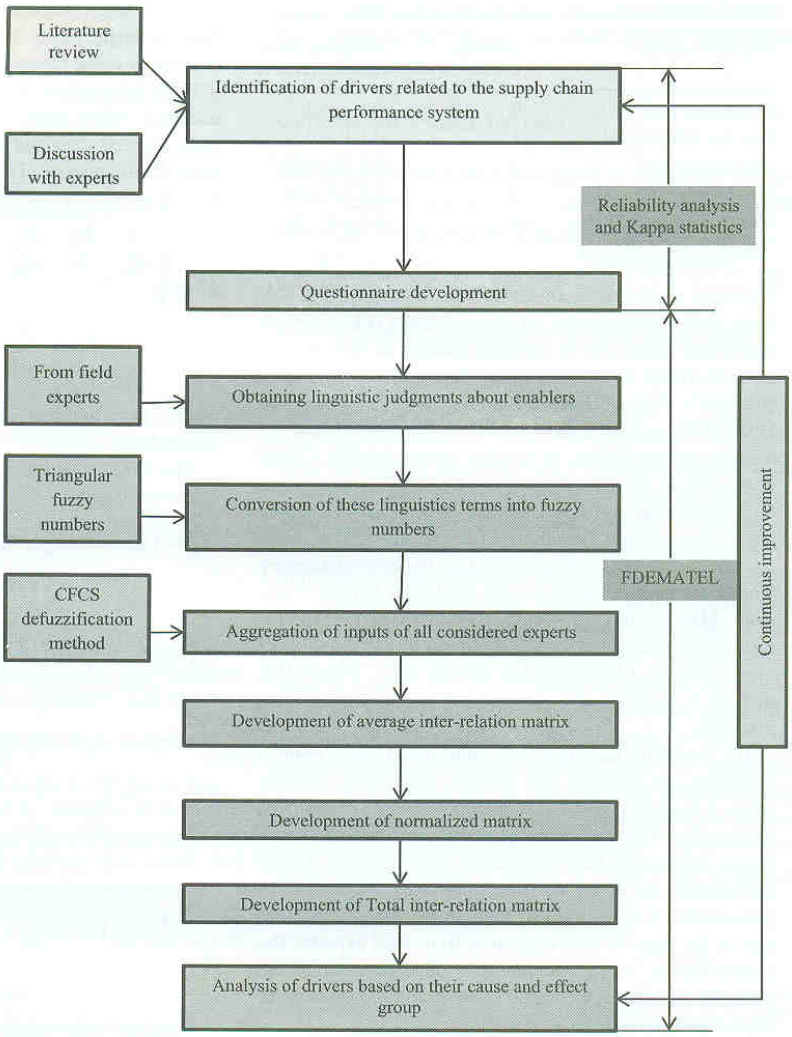


Fig. 1 Flow chart of the work

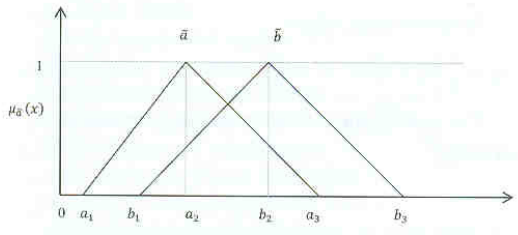


Fig. 2 Two triangular fuzzy numbers

Fuzzy DEMATEL Approach

The DEMATEL method (Gabus and Fontela 1973) analyze the mutual interaction among the complex set of criteria/factors on the basis of graph theory (Chen and Chen 2010) and the contextual relationship among the existing criteria, based on cause and effect behavior (Charan et al. 2008; Li and Tzeng 2009). In recent years, DEMATEL method has been applied by various scholars in different–different fields to solve the critical problems. Out of them some

Table 6 The fuzzy linguistic scale (modified from [35])

Linguistic variable	Influence score	Triangular fuzzy numbers (TFN's)
No influence (No)	0	(0, 0.1, 0.3)
Very low influence (VL)	1	(0.1, 0.3, 0.5)
Low influence (L)	2	(0.3, 0.5, 0.7)
High influence (H)	3	(0.5, 0.7, 0.9)
Very high influence (VH)	4	(0.7, 0.9, 1.0)

important applications are as follows: Hsu et al. (2007) have used a combination of fuzzy DEMATEL approach with fuzzy multi-criteria decision making to find the significant criteria in order to frame the relations of an ideal customer's choice behavior. Shieh et al. (2010) have used DEMATEL method in the hospital management to analyze the key success factors of hospital service quality. Chang et al. (2011) have applied fuzzy DEMATEL method to find significant criteria for the selection of supply chain management suppliers. Lin et al. (2011) have applied DEMATEL method to analyze the interrelationship among the core competences of IC design Service Company. Mavi et al. (2013) have applied fuzzy DEMATEL method to evaluate the logistical criteria for the selection of green supplier. Lin (2013) has applied fuzzy DEMATEL method to evaluate the green supply chain management practices.

By applying this method, both direct and indirect influences can be taken into account very easily. Furthermore, it has been seen that human perceptions or opinions on considered factors are subjective. In most of the cases, human judgements have uncertainty and vagueness in their preferences. To cover the uncertainty and vagueness of judgements, it is required to involve the fuzzy logic/concept at the stage of data collection from field experts. The criteria/factors in the questionnaire have been structured using fuzzy scale as shown in Table 2, and the opinions of field experts have been gathered.

Algorithm of fuzzy DEMATEL approach is given in five steps as:

(1) Construction of average direct-relation matrix.
(2) Normalization of direct relation matrix
(3) Determination of total-relation matrix
(4) Formation of a causal diagram
(5) Calculate threshold value

Construction of Average Direct-Relation Matrix

To collect the opinions of field expert's and find out the direct influence of i → jth element, fuzzy linguistic scale has been used as shown in Table 6. Then assign the corresponding triangular fuzzy numbers and aggregate the fuzzy numbers. Fuzzy aggregation process in this paper has been done using Converting Fuzzy data into Crisp Scores (CFCS) defuzzification method. The fuzzy maximum and minimum fuzzy number range have been determined and crisp score is obtained in the form of weighted average value (Opricovic and Tzeng 2003; Wu and Lee 2007). Let, $\tilde{A}_{ij}^{n} = \left(a_{ij}^{n}, b_{ij}^{n}, c_{ij}^{n}\right)$, and influence matrix is calculated as:

$$\tilde{A}^{n} = \begin{bmatrix} 0 & \tilde{A}_{12}^{n} & \tilde{A}_{13}^{n} & \dots & \tilde{A}_{1m}^{n} \\ \tilde{A}_{21}^{n} & 0 & \tilde{A}_{23}^{n} & \dots & \tilde{A}_{2m}^{n} \\ . & . & . & . & . \\ . & . & . & . & . \\ \tilde{A}_{m1}^{n} & \tilde{A}_{m2}^{n} & \tilde{A}_{m3}^{n} & \dots & 0 \end{bmatrix}, n = 1, 2, 3\ldots, h, \tag{10}$$

where m = number of criteria and h = number of respondents.

The CFCS method in five steps is given below:

Normalization:

$$xc_{ij}^{n} = \left(c_{ij}^{n} - \min a_{ij}^{n}\right) / \Delta_{min}^{max} \tag{11}$$

$$xb_{ij}^{n} = \left(b_{ij}^{n} - \min a_{ij}^{n}\right) / \Delta_{min}^{max} \tag{12}$$

$$xa_{ij}^{n} = \left(a_{ij}^{n} - \min a_{ij}^{n}\right) / \Delta_{min}^{max} \tag{13}$$

where $\Delta_{min}^{max} = \max c_{ij}^{n} - \min a_{ij}^{n}$

Compute right (rs) and left (ls) normalized values:

$$xcs_{ij}^{n} = xc_{ij}^{n} / \left(1 + xc_{ij}^{n} - xb_{ij}^{n}\right) \tag{14}$$

$$xas_{ij}^{n} = xb_{ij}^{n} / \left(1 + xb_{ij}^{n} - xa_{ij}^{n}\right) \tag{15}$$

Compute total normalized crisp values:

$$x_{ij}^{n} = \frac{\left[xas_{ij}^{n}\left(1 - xas_{ij}^{n}\right) + xcs_{ij}^{n} \times xcs_{ij}^{n}\right]}{\left[1 - xas_{ij}^{n} + xcs_{ij}^{n}\right]} \tag{16}$$

Compute crisp values:

$$z_{ij}^{n} = \min a_{ij}^{n} + x_{ij}^{n} \times \Delta_{min}^{max} \tag{17}$$

Integrate crisp values:

$$z_{ij} = \frac{1}{h}\left(z_{ij}^{1} + z_{ij}^{2} + \ldots + z_{ij}^{h}\right) \tag{18}$$

The average direct-relation matrix is written as:

$$Z = \left[z_{ij}\right] \tag{19}$$

Normalization of Direct Relation Matrix

Normalized direct relation matrix 'X' can be obtained by using the following equation:

$$X = \lambda \times Z, \tag{20}$$

where

$$\lambda = min\left[\frac{1}{\max\limits_{1 \le i \le m} \sum_{j=1}^{m} |z_{ij}|}, \frac{1}{\max\limits_{1 \le j \le m} \sum_{i=1}^{m} |z_{ij}|}\right],$$
$$i, j = 1, 2, \ldots, m$$

Determination of Total-Relation Matrix

It is written as:

$$T = X(I - X)^{-1}, \tag{21}$$

where 'I' is the identity matrix.

In the total relation matrix, the sum of rows (R) and columns (C) are calculated as:

$$T = [t_{ij}]_{m \times m}, \quad i, j = 1, 2 \ldots, m \tag{22}$$

$$R = \left[\sum_{j=1}^{m} t_{ij}\right]_{m \times 1} = [t_i]_{m \times 1} \tag{23}$$

$$C = \left[\sum_{i=1}^{m} t_{ij}\right]_{1 \times m} = [t_j]_{1 \times m} \tag{24}$$

Then, add the corresponding values of 'R' and 'C', to obtain the value of Prominence (R + C) and Relation (R − C).

Formation of a Causal Diagram

The casual diagram is prepared by mapping Prominence and Relation data, marked horizontally and vertically in the graph. Then categorization of enablers into cause and effect group is done. If (R − C) value is positive, enablers comes in the cause group and if (R − C) value is negative then comes in the category of effect group (Tzeng et al. 2007; Yang et al. 2008; Lee et al. 2009).

Calculate Threshold Value (α)

The threshold value (α) is calculated by taking the average of the elements in total relation matrix *T*, as in Eq. 25, and inter-relationship diagram (Fig. 4) is obtained.

$$\alpha = \frac{\sum_{i=1}^{m} \sum_{j=1}^{m} [t_{ij}]}{N}, \tag{25}$$

where N is the total number of elements in the matrix *T*.

Numerical Illustration

After performing reliability and consistency analysis of enablers, a questionnaire has been structured on fuzzy linguistic scale shown in Table 6, and discussed with nineteen field experts for seeking their opinion regarding enabler's importance. Then, linguistic tables are prepared; Table 9 is the example of such a table from one of the experts. This table explains the influencing nature of considered enablers over each other, for example the enabler A_1 has no influence on enabler A_{12}, low influence on enablers A_6, A_7, A_9, A_{10}, A_{11} and high influence over the enablers A_2, A_4, A_5, A_8, and very high influence on enabler A_3. After that linguistic scale direct-relation matrix (Table 9 in Appendix) has been converted into the corresponding triangular fuzzy numbers and then applied CFCS defuzzification method (Eqs. 11–17) to aggregate the fuzzy inputs. The defuzzified direct-relation matrix is shown in Table 10 in Appendix.

Table 7 Scores of each enabler and related values for cause and effect group

Enablers (A_m)	R	C	R + C	R − C
A_1. Higher customer satisfaction	3.7313	2.6959	6.4272	1.0354
A_2. Proper capacity utilization	1.7527	2.7207	4.4735	−0.9680
A_3. Making proper performance measurement	2.9198	2.5207	5.4405	0.3991
A_4. Information sharing	3.1850	2.7450	5.9300	0.4400
A_5. Effective integration of suppliers	2.5555	2.5568	5.1123	−0.0012
A_6. Higher degree of integration across the supply chain	3.6275	2.6255	6.2530	1.0021
A_7. Information availability and customer responsiveness	2.6021	2.9308	5.5329	−0.3287
A_8. Alignment of systems throughout the supply chain	3.3021	2.8167	6.1188	0.4854
A_9. Open and transparent communication	2.3929	2.5398	4.9327	−0.1470
A_{10}. Reduction of New Product Development cycle time	1.9673	3.5270	5.4943	−1.5598
A_{11}. Adoption of Modern manufacturing practices	2.1303	2.5282	4.6585	−0.3980
A_{12}. Effectiveness of scheduling techniques	2.8838	3.0198	5.9036	−0.1361
A_{13}. Long-term, strategic relationships	2.7165	2.8791	5.5956	−0.1626
A_{14}. Quality of the supply chain relationships	2.6877	2.8009	5.4886	−0.1133
A_{15}. Adoption of a market orientation	3.3430	2.8905	6.2335	0.4525

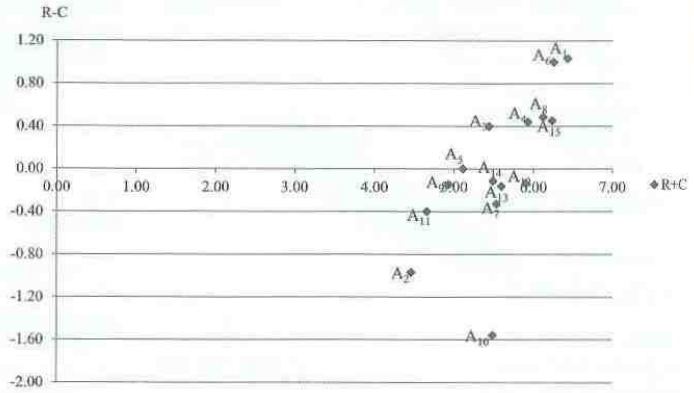


Fig. 3 Cause and effect diagram

Table 8 Interaction matrix of enablers ($\alpha = 0.1847$)

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1		*	*	*	*	*	*	*	*	*	*	*	*	*	*
A_2										*					
A_3	*			*			*	*	*	*		*	*	*	*
A_4	*				*					*		*	*	*	*
A_5				*						*					
A_6	*	*	*	*	*		*	*	*	*	*	*	*	*	*
A_7	*		*	*						*		*		*	*
A_8	*	*	*	*	*	*	*		*	*	*	*	*	*	*
A_9										*				*	
A_{10}												*			
A_{11}										*					
A_{12}	*	*					*	*		*			*	*	
A_{13}				*	*	*	*	*	*	*				*	*
A_{14}						*	*	*		*		*			*
A_{15}	*	*	*	*		*	*	*	*	*	*	*	*	*	

By using Eq. 18, aggregate the defuzzified/crisp direct-relation matrices of all experts and average direct-relation matrix is computed as given in Table 11 in Appendix.

Further, normalization of average direct-relation matrix has been done by using Eq. 20, as shown in Table 12 in Appendix. Then, total relation matrix as given in Table 13 in Appendix, is attained using Eq. 21. Using Eqs. 22–24, computed the sum of rows (R) and sum of columns (C) of the total inter-relation matrix to find the 'prominence' and 'relation' value of enablers and are given in Table 7. By mapping (R + C) and (R − C) values, cause and effect diagram is constructed as shown in Fig. 3 and threshold value ($\alpha = 0.1847$) by Eq. 25 has also been computed. The interaction matrix is constructed based on computed threshold value, to give a mutual degree of interaction among the considered enablers as shown in Table 8. Then an inter-relationship diagram is developed based on interaction matrix and shown in Fig. 4.

Results and Discussion

Table 7 and Fig. 5 show the importance ratings of all enablers based on their (R + C) values and noticed that the enabler A_1 has greater importance rating than others. The importance rating order of all enablers comes as: $A_1 > A_6 > A_{15} > A_8 > A_4 > A_{12} > A_{13} > A_7 > A_{10} > A_{14} > A_3 > A_5 > A_9 > A_{11} > A_2$. Thus, the primary focus of a company should be to fulfill the customer requirements to achieve better customer satisfaction. It is the role of enabler, 'higher degree of integration across supply chain (A_6)'.

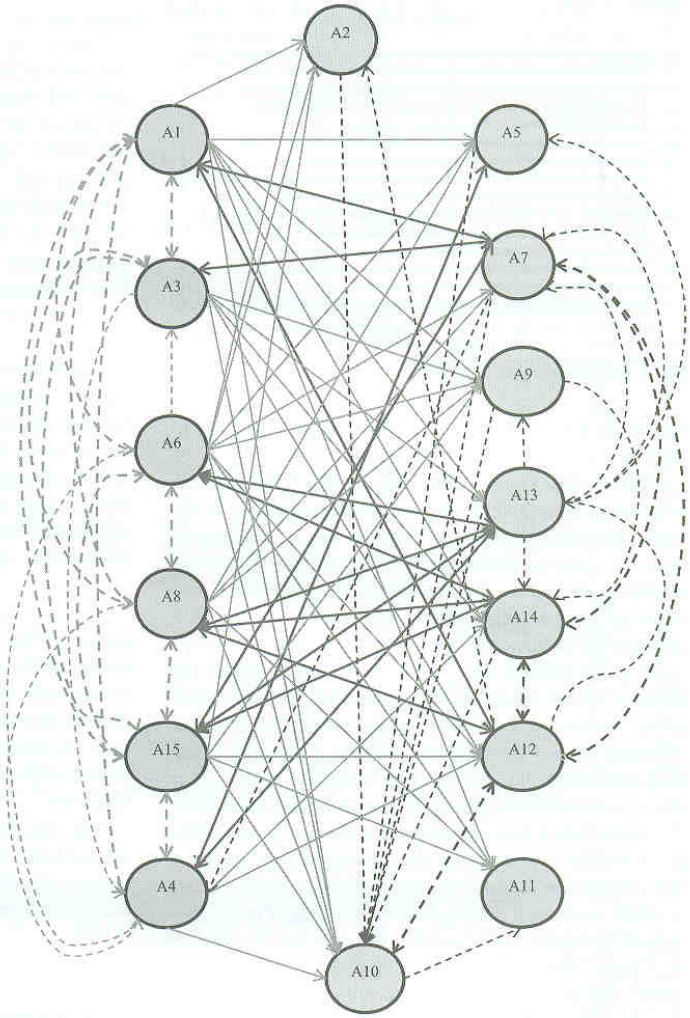


Fig. 4 Inter-relationship diagram

Also visualizing the (R − C) values, where enablers are categorized into their cause and effect group. The enablers namely higher customer satisfaction (A_1), making proper performance measurement (A_3), information sharing (A_4), higher degree of integration across the supply chain (A_6), alignment of systems throughout the supply chain (A_8) and adoption of a market orientation (A_{15}) have 'positive' (R − C) values and comes under the category of cause group. While the others such as: proper capacity utilization (A_2), effective integration of suppliers (A_5), information availability and customer responsiveness (A_7), open and transparent communication (A_9), reduction of new product development cycle time (A_{10}), adoption of modern manufacturing practices (A_{11}), effectiveness of scheduling techniques (A_{12}), long-term, strategic relationships (A_{13}), and quality of the supply chain relationships (A_{14}) have 'negative' (R − C) values and comes under the category of effect group. The enablers of cause group have an influence over the effect group enablers, which is shown by casual diagram (Fig. 3). Various thoughtful decisions may be made through Table 7 and by analyzing the casual diagram i.e. A1 seems to be more significant as it has highest (R + C) and (R − C) values i.e. 6.4272 and 1.0354 and also called precise influencing enabler, while the other enabler A_2 is considered as ample influencing enabler because it has



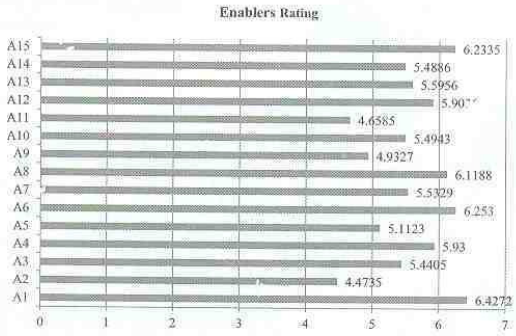


Fig. 5 Importance rating of enablers

lowest (R + C) and (R − C) values [4.4735 and minus 0.9680].

Table 8 and Fig. 4, show the mutual interaction or influencing behaviour of enablers over each other. In the inter-relationship diagram as shown in Fig. 4, six types of lines are used to represent an influencing nature of one enabler over another enabler based on their cause and effect group categorization as: thin regular orange line indicates the influence of one cause group enabler over the effect group enabler, thin dotted orange line indicates the influence of one cause group enabler over the other cause group enabler, thick dotted orange lines indicates the influence of two cause group enablers over each other, thick regular red line indicates the influence of one cause group enabler and one effect group enabler over each other, thin dotted black line indicates the influence of one effect group enabler over the other effect group enabler and thick dotted black line indicates the influence of two effect group enablers over each other. After visualizing Fig. 4, it is noticed that cause group enablers (A_1, A_3, A_4, A_6, A_8, and A_{15}) have higher degree of interaction and influencing nature over the effect group enablers (A_2, A_5, A_7, A_9, A_{10}, A_{11}, A_{12}, A_{13}, and A_{14}). Furthermore for better understanding can be says that enablers A_1, A_6, A_8, and A_{15} have highest degree of interaction and other enablers A_2, A_9, A_{10} and A_{11} have least degree of interaction, while the remaining enablers A_3, A_4, A_5, A_7, A_{12}, A_{13} and A_{14} have moderate degree of interaction.

Conclusions

To meet the specified research objective, a set of fifteen critical enablers is identified and analyzed using fuzzy DEMATEL approach. The proposed approach is an extension of DEMATEL method with use of linguistic terms and fuzzy aggregation method, in order to avoid the vagueness and impreciseness of the judgments. This study uses the CFCS (Converting Fuzzy data into Crisp Scores) method for fuzzy input aggregation. In order to reduce the complexity of problem and make precise decision, fuzzy DEMATEL categorizes a set of enablers into cause and effect groups.

The analysis reveals that the cause group enablers (A_1, A_3, A_4, A_6, A_8 and A_{15}) have more impact over the specified goal due to their higher degree of interaction. That's why; they cannot be moved easily and remunerated more attention. Based on research findings it is observed that, the higher customer satisfaction (A_1) is a very influencing enabler and can be achieved through adopting the market orientation (A_{15}). For this, a company should be focused towards higher degree of integration across the supply chain (A_6) in order to build an alignment of the systems throughout the supply chain (A_8). Hence, it is clear that enablers (A_1, A_6, A_8 and A_{15}) have higher degree of interaction through maintaining an appropriate level of information sharing (A_4) and proper performance measurement (A_3). On the other flank, the effect group enablers (A_2, A_5, A_7, A_9, A_{10}, A_{11}, A_{12}, A_{13} and A_{14}) are influenced easily by the cause group enablers and can be moved easily, hence, needs to be improved. Therefore, it is concluded that to improve or implement an effective and flexible SCPM system, an organization should emphasize more to maintain higher customer satisfaction level and higher degree of integration across their supply chain functions.

The concluding trail of the present research could be supported by the outcomes of previous studies. Some significant outcomes out of them are as: the higher level of customer satisfaction is an essential factor to enhance the market share value of an organization. The increase in the market share will provide a higher growth in the market place and improve the overall performance of a company (Anderson et al. 1994; Gunasekaran et al. 2001). Frohlich and Westbrook (2001) stated that if degree of integration among the supply chain stages is higher, then it is necessary that firm performance will also be better. Hence, it is easy to conclude that higher customer satisfaction and higher degree of integration across supply chain functions work as the key enablers for flexible SCPM system implementation.

The present study delivers an empirical analysis of considered enablers and highlights their categorization into cause and effect groups and degree of interaction. The outcomes of this study may provide an aid to the managers to implement an effective and flexible SCPM system through which overall profitability of an organization may be improved.

Limitations and Future Scope

During this study, respondents were selected from Indian automobile manufacturing industries located at National Capital Region, while the automobile industries are spread in a very large area of India at various locations with different requirements. Therefore; results of the present study cannot be generalized for automobile industries all over the country. This limitation generates future scope for this study as: first the data can be collected from other locations of India and compare the results. Second, other multi-criteria decision making techniques can also be applied in order to validate the results of this study.

Appendix

See Tables 9, 10, 11, 12, 13.

Table 9 The linguistic scale direct-relation matrix by expert 1

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1	No	H	VH	H	H	L	L	H	L	L	L	No	H	VL	VH
A_2	VL	No	VL	No	No	No	No	VL	No	VH	VL	L	VL	No	No
A_3	L	No	No	L	VL	VL	H	L	L	H	No	L	L	L	VH
A_4	No	L	VL	No	L	H	H	VL	H	L	VL	H	VH	H	H
A_5	VL	L	L	VH	No	L	No	L	VL	L	VL	VL	L	VL	VL
A_6	L	H	VL	VH	H	No	VH	VH	H	VH	L	L	H	H	L
A_7	H	No	L	H	No	No	No	No	No	L	VL	H	VL	L	H
A_8	VH	L	L	VL	VL	H	L	No	L	H	L	VH	VH	L	VL
A_9	No	No	No	L	L	L	L	VL	No	VL	No	No	L	VH	L
A_{10}	VL	H	No	No	L	No	No	No	No	No	VL	VH	No	No	No
A_{11}	L	No	No	No	No	No	No	L	No	VH	No	VL	VL	No	L
A_{12}	H	H	VL	VL	No	VL	H	H	VL	H	VL	No	L	H	VL
A_{13}	VL	No	No	L	L	L	L	L	H	L	L	No	No	L	L
A_{14}	L	No	No	VH	VL	H	H	VL	VH	H	VL	VL	H	No	H
A_{15}	VH	L	H	VL	No	VL	VH	L	VL	VH	H	H	VL	No	No

Table 10 Defuzzified direct-relation-matrix by expert 1

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1	0.0000	0.7622	0.8748	0.6905	0.7622	0.5524	0.5000	0.6905	0.5000	0.5000	0.5524	0.1252	0.6905	0.3095	0.8748
A_2	0.3095	0.0000	0.3095	0.1252	0.1402	0.1402	0.1252	0.3095	0.1252	0.8748	0.3427	0.5000	0.3095	0.1252	0.1252
A_3	0.5000	0.1402	0.0000	0.5000	0.3427	0.3427	0.6905	0.5000	0.5000	0.6905	0.1402	0.5000	0.5000	0.5000	0.8748
A_4	0.1252	0.5524	0.3095	0.0000	0.5524	0.7622	0.6905	0.3095	0.6905	0.5000	0.3427	0.6905	0.8748	0.6905	0.6905
A_5	0.3095	0.5524	0.5000	0.8748	0.0000	0.5524	0.1252	0.5000	0.3095	0.5000	0.3427	0.3095	0.5000	0.3095	0.3095
A_6	0.5000	0.7622	0.3095	0.8748	0.7622	0.0000	0.8748	0.8748	0.6905	0.8748	0.5524	0.5000	0.6905	0.6905	0.5000
A_7	0.6905	0.1402	0.5000	0.6905	0.1402	0.1402	0.0000	0.1252	0.1252	0.5000	0.3427	0.6905	0.3095	0.5000	0.6905
A_8	0.8748	0.5524	0.5000	0.3095	0.3427	0.7622	0.5000	0.0000	0.5000	0.6905	0.5524	0.8748	0.8748	0.5000	0.3095
A_9	0.1252	0.1402	0.1252	0.5000	0.5524	0.5524	0.5000	0.3095	0.0000	0.3095	0.1402	0.1252	0.5000	0.8748	0.5000
A_{10}	0.3095	0.7622	0.1252	0.1252	0.5524	0.1402	0.1252	0.1252	0.1252	0.0000	0.3427	0.8748	0.1252	0.1252	0.1252
A_{11}	0.5000	0.1402	0.1252	0.1252	0.1402	0.1402	0.1252	0.5000	0.1252	0.8748	0.0000	0.3095	0.3095	0.1252	0.5000
A_{12}	0.6905	0.7622	0.3095	0.3095	0.1402	0.3427	0.6905	0.6905	0.3095	0.6905	0.3427	0.0000	0.5000	0.6905	0.3095
A_{13}	0.3095	0.1402	0.1252	0.5000	0.5524	0.5524	0.5000	0.5000	0.6905	0.5000	0.5524	0.1252	0.0000	0.5000	0.5000
A_{14}	0.5000	0.1402	0.3095	0.8748	0.3427	0.7622	0.6905	0.3095	0.8748	0.6905	0.3427	0.3095	0.6905	0.0000	0.6905
A_{15}	0.8748	0.5524	0.6905	0.3095	0.1402	0.3427	0.8748	0.5000	0.3095	0.8748	0.7622	0.6905	0.3095	0.1252	0.0000

Table 11 Average direct-relation matrix

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1	0.0000	0.7390	0.7619	0.7872	0.7357	0.5882	0.5920	0.7827	0.6967	0.5799	0.6260	0.3632	0.7494	0.4554	0.7929
A_2	0.4376	0.0000	0.3171	0.1646	0.1641	0.1838	0.1503	0.3631	0.1663	0.8159	0.3523	0.5864	0.3354	0.1365	0.1284
A_3	0.5429	0.2059	0.0000	0.5100	0.3668	0.3466	0.6704	0.6054	0.5266	0.6692	0.1690	0.5753	0.4872	0.5964	0.7518
A_4	0.1816	0.5992	0.3597	0.0000	0.5397	0.7097	0.7010	0.3134	0.7068	0.5100	0.3865	0.6968	0.7988	0.7456	0.6725
A_5	0.3292	0.5373	0.5780	0.7969	0.0000	0.5343	0.1595	0.5156	0.3745	0.5501	0.3428	0.3735	0.4882	0.3040	0.3165
A_6	0.5517	0.7200	0.3770	0.7289	0.7767	0.0000	0.7404	0.8042	0.6674	0.8651	0.5570	0.5356	0.7012	0.7065	0.5716
A_7	0.6813	0.2062	0.6509	0.6795	0.1929	0.2148	0.0000	0.1753	0.1462	0.5401	0.3938	0.7265	0.3057	0.5356	0.7333
A_8	0.7339	0.6196	0.5490	0.3607	0.4043	0.6900	0.5712	0.0000	0.5155	0.6698	0.5953	0.8047	0.7691	0.5466	0.3265
A_9	0.1756	0.2282	0.1626	0.5218	0.6366	0.5530	0.5295	0.3735	0.0000	0.3402	0.1794	0.1759	0.5183	0.8335	0.5211
A_{10}	0.3161	0.6711	0.3694	0.1644	0.6018	0.2375	0.2186	0.2239	0.3664	0.0000	0.3441	0.7850	0.2556	0.1953	0.1587
A_{11}	0.5763	0.2262	0.2125	0.2710	0.2749	0.2933	0.2320	0.5466	0.1659	0.7571	0.0000	0.3334	0.3060	0.2845	0.5933
A_{12}	0.7025	0.7112	0.3968	0.3399	0.2042	0.3586	0.7100	0.7048	0.3534	0.6692	0.4487	0.0000	0.5383	0.7075	0.2858
A_{13}	0.3029	0.2356	0.2442	0.5301	0.6095	0.5615	0.5095	0.5957	0.7061	0.5297	0.5755	0.1853	0.0000	0.5055	0.5421
A_{14}	0.5535	0.2660	0.4436	0.8163	0.4109	0.7191	0.6722	0.3227	0.8435	0.6598	0.3542	0.4428	0.7096	0.0000	0.7705
A_{15}	0.7305	0.6276	0.7432	0.5195	0.3074	0.4502	0.7396	0.6251	0.4845	0.7678	0.7931	0.7255	0.4364	0.3464	0.0000

Table 12 Normalized direct-relation matrix

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1	0.0000	0.0799	0.0824	0.0851	0.0795	0.0636	0.0640	0.0846	0.0753	0.0627	0.0677	0.0393	0.0810	0.0492	0.0857
A_2	0.0473	0.0000	0.0343	0.0178	0.0177	0.0199	0.0163	0.0393	0.0180	0.0882	0.0381	0.0634	0.0363	0.0148	0.0139
A_3	0.0587	0.0223	0.0000	0.0551	0.0397	0.0375	0.0725	0.0655	0.0569	0.0723	0.0183	0.0622	0.0527	0.0645	0.0813
A_4	0.0196	0.0648	0.0389	0.0000	0.0583	0.0767	0.0758	0.0339	0.0764	0.0551	0.0418	0.0753	0.0864	0.0806	0.0727
A_5	0.0356	0.0581	0.0625	0.0861	0.0000	0.0578	0.0172	0.0557	0.0405	0.0595	0.0371	0.0404	0.0528	0.0329	0.0342
A_6	0.0596	0.0778	0.0346	0.0788	0.0840	0.0000	0.0800	0.0869	0.0721	0.0724	0.0602	0.0579	0.0758	0.0764	0.0618
A_7	0.0737	0.0223	0.0704	0.0735	0.0209	0.0232	0.0000	0.0190	0.0158	0.0584	0.0426	0.0785	0.0330	0.0579	0.0793
A_8	0.0793	0.0670	0.0593	0.0390	0.0437	0.0746	0.0618	0.0000	0.0557	0.0724	0.0644	0.0870	0.0831	0.0591	0.0353
A_9	0.0190	0.0247	0.0176	0.0564	0.0688	0.0598	0.0572	0.0404	0.0000	0.0368	0.0194	0.0190	0.0560	0.0901	0.0563
A_{10}	0.0342	0.0725	0.0399	0.0178	0.0651	0.0257	0.0236	0.0242	0.0396	0.0000	0.0372	0.0849	0.0276	0.0211	0.0172
A_{11}	0.0623	0.0245	0.0230	0.0293	0.0297	0.0317	0.0251	0.0591	0.0179	0.0818	0.0000	0.0360	0.0331	0.0308	0.0641
A_{12}	0.0759	0.0769	0.0429	0.0368	0.0221	0.0388	0.0768	0.0762	0.0382	0.0723	0.0485	0.0000	0.0582	0.0765	0.0309
A_{13}	0.0327	0.0255	0.0264	0.0573	0.0659	0.0607	0.0551	0.0644	0.0763	0.0573	0.0622	0.0200	0.0000	0.0547	0.0586
A_{14}	0.0324	0.0259	0.0476	0.0443	0.0444	0.0777	0.0727	0.0349	0.0346	0.0713	0.0383	0.0479	0.0443	0.0000	0.0833
A_{15}	0.0790	0.0678	0.0803	0.0562	0.0332	0.0487	0.0800	0.0676	0.0524	0.0830	0.0857	0.0784	0.0472	0.0375	0.0000

Table 13 Total inter-relation matrix

	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}
A_1	0.1798	0.2558	0.2453	0.2629	0.2454	0.2359	0.2556	0.2667	0.2413	0.2947	0.2327	0.2404	0.2685	0.2339	0.2726
A_2	0.1303	0.0878	0.1117	0.1011	0.0984	0.1003	0.1067	0.1264	0.0974	0.1934	0.1159	0.1556	0.1244	0.1013	0.1020
A_3	0.1974	0.1657	0.1342	0.1965	0.1722	0.1746	0.2239	0.2081	0.1877	0.2514	0.1521	0.2186	0.2009	0.2085	0.2280
A_4	0.1732	0.2148	0.1799	0.1567	0.2006	0.2213	0.2390	0.1928	0.2156	0.2531	0.1843	0.2418	0.2432	0.2362	0.2325
A_5	0.1562	0.1809	0.1728	0.2051	0.1184	0.1758	0.1521	0.1817	0.1573	0.2175	0.1508	0.1780	0.1836	0.1610	0.1644
A_6	0.2305	0.2496	0.1975	0.2517	0.2443	0.1716	0.2636	0.2620	0.2321	0.2963	0.2214	0.2513	0.2580	0.2528	0.2448
A_7	0.1962	0.1508	0.1863	0.1966	0.1392	0.1454	0.1400	0.1516	0.1358	0.2203	0.1598	0.2163	0.1666	0.1858	0.2114
A_8	0.2340	0.2243	0.2043	0.1996	0.1939	0.2242	0.2314	0.1680	0.2032	0.2765	0.2104	0.2583	0.2476	0.2216	0.2050
A_9	0.1331	0.1403	0.1266	0.1733	0.1756	0.1717	0.1811	0.1579	0.1095	0.1860	0.1285	0.1489	0.1766	0.2050	0.1788
A_{10}	0.1284	0.1653	0.1263	0.1133	0.1496	0.1157	0.1241	0.1240	0.1255	0.1250	0.1234	0.1848	0.1276	0.1186	0.1156
A_{11}	0.1631	0.1313	0.1209	0.1323	0.1275	0.1308	0.1362	0.1640	0.1158	0.2117	0.0987	0.1522	0.1420	0.1346	0.1691
A_{12}	0.2112	0.2123	0.1721	0.1756	0.1536	0.1723	0.2224	0.2153	0.1668	0.2505	0.1772	0.1573	0.2029	0.2147	0.1795
A_{13}	0.1620	0.1571	0.1482	0.1888	0.1881	0.1867	0.1944	0.1965	0.1954	0.2247	0.1820	0.1672	0.1403	0.1888	0.1964
A_{14}	0.1631	0.1579	0.1682	0.1756	0.1658	0.1981	0.2108	0.1695	0.1559	0.2369	0.1599	0.1934	0.1802	0.1349	0.2174
A_{15}	0.2374	0.2270	0.2265	0.2157	0.1841	0.2011	0.2495	0.2322	0.2006	0.2891	0.2312	0.2557	0.2166	0.2032	0.1729

References

Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing, 58*(7), 53–66.

Anderson, M. F., & Katz, P. B. (1998). Strategic sourcing. *International Journal of Logistics Management, 9*(1), 1–13.

Aviv, Y. (2007). On the benefits of collaborative forecasting partnerships between retailers and manufacturers. *Management Science, 53*(5), 777–794.

Barut, M., Faisst, W., & Kanet, J. J. (2002). Measuring supply chain coupling: an information system perspective. *European Journal of Purchasing & Supply Management, 8*(3), 161–171.

Beamon, B. M. (1999). Measuring supply chain performance. *International Journal of Operations and Production Management, 19*(3), 275–292.

Bhagwat, R., & Sharma, M. K. (2007). Performance measurement of supply chain management: A balanced scorecard approach. *Computers & Industrial Engineering, 53*(1), 43–62.

Bowersox, D., Closs, D., & Cooper, M. (2010). *Supply chain logistics management* (Int ed.). New York: McGraw-Hill.

Brown, S., Squire, B., & Blackmon, K. (2007). The contribution of manufacturing strategy involvement and alignment to world-class manufacturing performance. *International Journal of Operations and Production Management, 27*(3), 282–302.

Chan, F. T. S., Nayak, A., Raj, R., Chong, A. Y. L., & Tiwari, M. (2014). An innovative supply chain performance measurement system incorporating research and development (R&D) and marketing policy. *Computers & Industrial Engineering, 69*, 64–70.

Chang, B., Chang, C. W., & Wu, C. H. (2011). Fuzzy DEMATEL method for developing supplier selection criteria. *Expert Systems with Applications, 38*(3), 1850–1858.

Charan, P., Shankar, R., & Baisya, R. K. (2008). Analysis of interactions among the variables of supply chain performance measurement system implementation. *Business Process Management Journal, 14*(4), 512–529.

Chen, J. K., & Chen, S. (2010). Using a novel conjunctive MCDM approach based on DEMATEL, fuzzy ANP, and TOPSIS as an innovation support system for Taiwanese higher education. *Expert Systems with Applications, 37*(3), 1981–1990.

Chiadamrong, N., & Prasertwattana, K. (2006). A comparative study of supply chain models under the traditional centralized and coordinating policies with incentive schemes. *Computers & Industrial Engineering, 50*(4), 367.

Cho, D. W., Lee, Y. H., Ahn, S. H., & Hwang, M. K. (2012). A framework for measuring the performance of service supply chain management. *Computers & Industrial Engineering, 62*(3), 801–818.

Christopher, M. (1998). Logistics and supply chain management: Strategies for reducing costs and improving services. *Financial times* (2nd ed.). London: Pitman Publishing.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin, 70*(4), 213–220.

Cooper, R. G., & Kleinschmidt, E. J. (1995). Benchmarking the firm's critical success factors in new project development. *Journal of Product Innovation Management, 12*(5), 374–391.

Cousins, P. D. (2005). The alignment of appropriate firm and supply strategies for competitive advantage. *International Journal of Operations & Production Management, 25*(5), 403–428.

Das, A., Narasimhan, R., & Talluri, S. (2006). Supplier integration: Finding an optimal configuration. *Journal of Operations Management, 24*(5), 563–582.

Daugherty, P. J., Ellinger, A. E., & Rogers, D. S. (1995). Information accessibility: Customer responsiveness and enhanced performance. *International Journal of Physical Distribution and Logistics Management, 25*(1), 4–17.

Dawes, J. (1999). The relationship between subjective and objective company performance measures in market orientation research: Further empirical evidence. *Marketing Bulletin, 10*, 65–75.

Demirbag, M., Koh, S. C. L., Tatoglu, E., & Zaim, S. (2006). TQM and market orientation's impact on SMEs' performance. *Industrial Management & Data Systems, 106*(8), 1206–1228.

Dreyer, D. E. (2000). Performance measurement: A practitioner's perspective. *Supply Chain Management Review, 4*(4), 30–36.

Fawcett, S. E., & Cooper, M. B. (1998). Logistics performance measurement and customer success. *Industrial Marketing Management, 27*(4), 341–357.

Fawcett, S. E., & Myers, M. B. (2001). Product and employee development in advanced manufacturing: Implementation and impact. *International Journal of Production Research, 39*(1), 65–79.

Flynn, B. B., & Flynn, E. J. (2005). Synergies between supply chain management and quality management: Emerging implications. *International Journal of Production Research, 43*(16), 3421–3436.

Frohlich, M. T., & Westbrook, R. (2001). Arcs of integration: an international study of supply chain strategies. *Journal of Operations Management, 19*(2), 185–200.

Fynes, B., de Búrca, S., & Marshall, D. (2004). Environmental uncertainty, supply chain relationship quality and performance. *Journal of Purchasing and Supply Management, 10*(4/5), 179–190.

Gabus, A., & Fontela, E. (1973). *Perceptions of the world problematique: Communication procedure, communicating with those bearing collective responsibility*. Geneva: Battelle Geneva Research Center. (DEMATEL Report No. 1).

Giannakis, M. (2011). Management of service supply chains with a service oriented reference model: The case of management consulting source. *Supply Chain Management, 16*(5), 346–361.

Graham, T. S., Dougherty, P. J., & Dudley, W. N. (1994). The long term strategic impact of purchasing partnerships. *International Journal of Purchasing and Materials Management, 30*(4), 13–18.

Green, K. W, Jr., & Inman, R. A. (2005). Using a just-in-time selling strategy to strengthen supply chain linkages. *International Journal of Production Research, 43*(16), 3437–3453.

Green, K. W., Jr., Inman, R. A., Brown, G., & Willis, T. H. (2005). Market orientation: Relation to structure and performance. *Journal of Business & Industrial Marketing, 20*(6), 276–284.

Griffin, A. (1997). PDMA research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management, 14*(6), 429–458.

Gunasekaran, A., & Kobu, B. (2007). Performance measures and metrics in logistics and supply chain management: A review of recent literature (1995–2004) for research and applications. *International Journal of Production Research, 45*(12), 2819–2840.

Gunasekaran, A., & Ngai, E. W. T. (2008). Adoption of e-procurement in Hong Kong: An empirical research. *International Journal of Production Economics, 113*(1), 159–175.

Gunasekaran, A., Patel, C., & McGaughey, R. E. (2004). A Framework for supply chain performance measurement. *International Journal of Production Economics, 87*(3), 333–347.

Gunasekaran, A., Patel, C., & Tirtiroglu, E. (2001). Performance measures and metrics in a supply chain environment. *International Journal of Operations & Production Management, 21*(1/2), 71–87.

Gupta, A. K., & Govindarajan, V. (2000). Knowledge flows within multinational corporations. *Strategic Management Journal, 21*(4), 473–496.

Handfield, R. B., & Nichols, E. L, Jr. (1999). *Introduction to supply chain management*. Prentice-Hall: Upper Saddle River.

Harrtman, R. V., & Bengtsson, L. (2009). Manufacturing competence: A key to success-full supplier integration. *International Journal of Manufacturing Technology and Management, 16*(3), 283–299.

Heizer, J., & Render, B. (2006). *Operations management* (8th ed.). Pearson Education New Jersey: Upper Saddle River.

Hofmann, E., & Locker, A. (2009). Value-based performance measurement in supply chains: A case study from the packaging industry. *Production Planning and Control, 20*(1), 68–81.

Holweg, M., Disney, S., Holmstrom, J., & Smaros, J. (2005). Supply chain collaboration: Making sense of the strategy continuum. *European Management Journal, 23*(2), 170–181.

Hsu, C. Y., Chen, Y. H., & Tzeng, G. H. (2007). FMCDM with Fuzzy DEMATEL Approach for customers' choice behavior model. *International Journal of Fuzzy Systems, 9*(4), 236–246.

Hu, F., Lim, C. C., & Lu, Z. (2013). Coordination of supply chains with a flexible ordering policy under yield and demand uncertainty. *International Journal of Production Economics, 146*, 686–693.

Hurley, R. F., & Hult, G. T. M. (1998). Innovation, market orientation, and organisational learning: An integration and empirical examination. *Journal of Marketing, 62*, 42–54.

Ittner, C. D., & Larcker, D. F. (1997). Product development cycle time and organizational performance. *Journal of Marketing Research, 34*(1), 13–23.

Jakhar, S. K., & Barua, M. K. (2014). An integrated model of supply chain performance evaluation and decision-making using structural equation modelling and fuzzy AHP. *Production Planning and Control, 25*(11), 938–957.

Johnson, J. L. (1999). Strategic integration in distribution channels: Managing the inter-firm relationship as a strategic asset. *Academy of Marketing Science Journal, 27*(1), 4–18.

Kaufmann, A., & Gupta, M. M. (1991). *Introduction to fuzzy arithmetic: theory and application*. New York: Van Nostrand Reinhold.

Kaynak, H. (2003). The relationship between total quality management practices and their effects on firm performance. *Journal of Operations Management, 21*(4), 405–435.

Ketokivi, M. A., & Schroeder, R. G. (2004). Strategic, structural contingency and institutional explanations in the adoption of manufacturing practices. *Journal of Operations Management, 22*(1), 63–89.

Klein, S., & Roth, V. J. (1993). Satisfaction with international marketing channels. *Journal of the Academy of Marketing Science, 21*(1), 39–44.

Koh, S. C. L., Saad, S. M., & Arunachalam, S. (2006). Competing in the 21st century supply chain through supply chain management and enterprise resource planning integration. *International Journal of Physical Distribution and Logistics Management, 36*(6), 455–465.

Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: The construct, research propositions, and managerial implications. *Journal of Marketing, 54*(2), 1–18.

Kumar, P., & Deshmukh, S. G. (2006). A model for flexible supply chain through flexible manufacturing. *Global Journal of Flexible Systems Management, 7*(3), 17–24.

Kwon, I.-W. G., & Suh, T. (2004). Factors affecting the level of trust and commitment in supply chain relationships. *Journal of Supply Chain Management, 40*(2), 4–14.

Labahn, D. W., Ali, A., & Krapfel, R. (1997). New product development cycle time: The influence of project and process factors in small manufacturing companies. *Journal of Business Research, 36*(2), 179–188.

Lalonde, B. J. (1998). Building a supply chain relationship. *Supply Chain Management Review, 2*(2), 7–8.

Landis, J. R., & Koch, G. G. (1977). A one-way component of variance model for categorical data. *Biometrics, 33*(4), 671–679.

Lawson, B., & Potter, A. (2012). Determinants of knowledge transfer in inter-firm new product development projects. *International Journal of Operations & Production Management, 32*(10), 1228–1247.

Lee, H. I., Kang, H. Y., Hsu, C. F., & Hung, H. C. (2009). A green supplier selection model for high-tech industry. *Expert System Application, 36*(4), 7917–7927.

Li, C. W., & Tzeng, G. H. (2009). Identification of a threshold value for the DEMATEL method using the maximum mean de-entropy algorithm to find critical services provided by a semiconductor intellectual property mall. *Expert Systems with Applications, 8*(1), 9891–9898.

Lin, R. J. (2013). Using fuzzy DEMATEL to evaluate the green supply chain management practices. *Journal of Cleaner Production, 40*, 32–39.



Lin, H. T., & Chang, W. L. (2008). Order selection and pricing methods using flexible quantity and fuzzy approach for buyer evaluation. *European Journal of Operational Research, 187*(2), 415–428.

Lin, C., & Tseng, H. (2006). Identifying the pivotal role of participation strategies and information technology application for supply chain excellence. *Industrial Management & Data Systems, 106*(5), 739–756.

Lin, Y. T., Yang, Y. H., Kang, J. S., & Yu, H. C. (2011). Using DEMATEL method to explore the core competences and causal effect of the IC design service company: An empirical case study. *Expert Systems with Applications, 38*(5), 6262–6268.

Lings, I. (2004). Internal market orientation construct and consequences. *Journal of Business Research, 57*(4), 405–413.

Little, D., Kenworthy, J., Jarvis, P., & Porter, K. (1995). Scheduling across the supply chain. *Logistics Information Management, 8*(1), 42–48.

Lockstrom, M., Schadel, J., Harrison, N., Moser, R., & Malhotra, M. K. (2010). Antecedents to supplier integration in the automotive, industry: A multiple-case study of foreign subsidiaries in China. *Journal of Operations Management, 28*(3), 240–256.

Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: From constructs to theory. *Journal of Operations Management, 16*(4), 407–425.

Mangla, S. K., Kumar, P., & Barua, M. K. (2014). Flexible decision approach for analyzing performance of sustainable supply chains under risks/uncertainty. *Global Journal of Flexible Systems Management, 15*(2), 113–130.

Mavi, R. K., Kazemi, S., Najafabadi, A. F., & Mousaabadi, H. B. (2013). Identification and assessment of logistical factors to evaluate a green supplier using the fuzzy logic DEMATEL method. *Polish journal of environmental studies, 22*(2), 445–455.

Mishra, A. A., & Shah, R. (2009). In union lies strength: Collaborative competence in new product development and its performance effects. *Journal of Operations Management, 27*(4), 324–338.

Morgan, J., & Monczka, R. M. (1996). Supplier integration: A new level of supply chain management. *Purchasing, 120*, 110–113.

Narasimhan, R., & Jayaram, J. (1998). Causal linkages in supply chain management: an exploratory study of North American manufacturing firms. *Decision Sciences, 29*(3), 579–605.

Neely, A., Gregory, M. J., & Platts, K. (1995). Performance measurement system design, a literature review and research agenda. *International Journal of Operations and Production Management, 15*(4), 80–116.

Nyaga, G., Whipple, J., & Lynch, D. (2010). Examining supply chain relationships: Do buyer and supplier perspectives on collaborative relationships differ? *Journal of Operations Management, 28*(2), 101–114.

Oke, A., Prajogo, D. I., & Jayaram, J. (2013). Strengthening the innovation chain: The role of internal innovation climate and strategic relationships with supply chain partners. *Journal of Supply Chain Management, 49*(4), 43–58.

Opricovic, S., & Tzeng, G. H. (2003). Defuzzification within a multicriteria decision model. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 11*(5), 635–652.

Paulraj, A., Lado, A. A., & Chen, J. J. (2008). Inter-organizational communication as a relational competency: Antecedents and performance outcomes in collaborative buyer-supplier relationships. *Journal of Operations Management, 26*(1), 45–64.

Petersen, K., Handfield, R., & Ragatz, G. (2005). Supplier integration into new product development: coordinating product, process, and supply chain design. *Journal of Operations Management, 23*(3–4), 371–388.

Qureshi, M. N., Kumar, D., & Kumar, P. (2008). An integrated model to identify and classify the key criteria and their role in the assessment of 3PL services providers. *Asia Pacific Journal of Marketing and Logistics, 20*(2), 227–249.

Rabelo, L., Eskandari, H., Shaalan, T., & Helal, M. (2007). Value chain analysis using hybrid simulation and AHP. *International Journal of Production Economics, 105*, 536–547.

Ragatz, G. L., Handfield, R. B., & Scannell, T. V. (1997). Success factors for integrating suppliers into new product development. *Journal of Product Innovation Management, 14*(3), 190–202.

Ravi, V., Shankar, R., & Tiwari, M. K. (2005). Analyzing alternatives in reverse logistics for end-of-life computers: ANP and balanced scorecard approach. *Computers & Industrial Engineering, 48*(2), 327–356.

Saad, M., & Patel, B. (2006). An investigation of supply chain performance measurement in Indian automotive sector. *Benchmarking, 13*(1/2), 36–53.

Shieh, J. I., Wu, H. H., & Huang, K. K. (2010). A DEMATEL method in identifying key success factors of hospital service quality. *Knowledge-Based Systems, 23*(3), 277–282.

Sidola, A., Kumar, P., & Kumar, D. (2012). System dynamics investigation of information technology in small and medium enterprise supply chain. *Journal of Advances in Management Research, 9*(2), 199–207.

Simatupang, T. M., & Sridharan, R. (2002). The collaborative supply chain. *International Journal of Logistics Management, 13*(1), 15–30.

Slack, N., Chambers, S., Harland, C., Harrison, A., & Johnston, R. (1995). *Operations management*. London: Pitman Publishing.

Sushil. (2012). Multiple perspectives of flexible systems management. *Global Journal of Flexible Systems Management, 13*(1), 1–2.

Thakkar, J., Kanda, A., & Deshmukh, S. G. (2009). Supply chain performance measurement framework for small and medium scale enterprises. *Benchmarking, 16*(5), 702–723.

Thome, A. M. T., Scavarda, L. F., Pires, S. R. I., Ceryno, P., & Klingebiel, K. (2014). A multi-tier study on supply chain flexibility in the automotive industry. *International Journal of Production Economics, 158*, 91–105.

Tompkins, J., & Ang, D. (1999). What are your greatest challenges related to supply chain performance measurement? *IIE Solutions, 31*(6), 66.

Turner, J. R. (1993). Integrated supply chain management: What's wrong with this picture? *Industrial Engineering, 25*(12), 52–55.

Tyagi, M., Kumar, P., & Kumar, D. (2014). Selecting alternatives for improvement in IT enabled supply chain performance. *International Journal of Procurement Management, 7*(2), 168–182.

Tzeng, G. H., Chiang, C. H., & Li, C. W. (2007). Evaluating intertwined effects in e-learning programs: a novel hybrid MCDM model based on factor analysis and DEMATEL. *Expert Systems with Applications, 32*(4), 1028–1044.

Ulaga, W., & Chacour, S. (2001). Measuring customer-perceived value in business markets. *Industrial Marketing Management, 30*(6), 525–540.

Van der Vorst, J. G. A. J., & Beulens, A. J. M. (2002). Identifying sources of uncertainty to generate supply chain redesign strategies. *International Journal of Physical Distribution & Logistics Management, 32*(6), 409–430.

Vanderhaeghe, A., & de Treville, S. (2003). How to fail at flexibility. *Supply Chain Forum, 4*(1), 64–67.

Wadhwa, S., Madaan, J., & Avneet, S. (2007). Need for flexibility and innovation in healthcare management systems. *Global Journal of Flexible Systems Management, 8*(1–2), 45–54.

Waggoner, D., Neely, A., & Kennerley, M. (1999). The forces that shape organizational performance measurement systems: An interdisciplinary review. *International Journal of Production Economics, 60–61*, 53–60.

Womack, J. P., & Jones, D. T. (2005). *Lean solutions: How companies and customers can create wealth together*. New York: Simon & Schuster.

Wu, W. W., & Lee, Y. T. (2007). Developing global managers' competencies using the fuzzy DEMATEL method. *Expert Systems with Applications, 32*(2), 499–507.

Yang, Y. P., Shieh, H. M., Leu, J. D., & Tzeng, G. H. (2008). A novel hybrid MCDM model combined with DEMATEL and ANP with applications. *International Journal Operational Research, 5*(3), 160–168.

Yeh, C. H., & Deng, H. (2004). A practical approach to fuzzy utilities comparison in fuzzy multi-criteria analysis. *International Journal of Approximate Reasoning, 35*(2), 179–194.

Zadeh, L. A. (1965). Fuzzy Set. *Information and Control, 8*(3), 338–353.

Key Questions

1. How to identify the significant enablers related to flexible supply chain performance measurement (SCPM) system?
2. How does the enabler of SCPM system can be used flexibly in improving the performance of an organization as a whole?
3. How does to address the internal consistency of enablers with general perspectives of an organization?
4. How do these enablers may useful for managers to improve and implement a flexible SCPM system?



Mohit Tyagi is research scholar in Mechanical and Industrial Engineering Department, IIT Roorkee, India and working in the field of Supply Chain Management, Corporate Social Responsibility, Flexible Systems, Modeling and Simulation of Supply Chain, Product Development. He has obtained his bachelor's degree (Hons) in Mechanical Engineering from Uttar Pradesh Technical University, Lucknow, India in 2008 and masters in Mechanical Engineering with specialization in Product Design and Development (Gold Medalist) from Motilal Nehru National Institute of Technology, Allahabad, India in 2010. He has published/presented several papers in repute journals and conferences.



Pradeep Kumar is a Professor in the department of Mechanical and Industrial Engineering at IIT Roorkee, India. He has obtained his bachelor's degree in industrial engineering in 1982 from Roorkee University and masters in production engineering (Gold Medalist) from University of Roorkee (Now, IIT Roorkee) in 1989. He has received Ph.D. degree in 1994 with specialization in Industrial Engineering from University of Roorkee (Now, IIT Roorkee), India. He has more than 25 years research/teaching experience. He has supervised 27 Ph.D. Thesis and had contributed more than 450 publications in international/national journals/conferences of repute. His research interests include Advanced Manufacturing Processes; Microwave Joining of Metals, Metal Casting; Industrial Engineering; Supply Chain Management (SCM), Quality Engineering: Robust Design Methodologies, Reliability Engineering.



Dinesh Kumar is Professor in Mechanical and Industrial Engineering Department, IIT Roorkee, India. He obtained his B.Sc. Engineering with Hons. (Mechanical engineering) in 1980 from Punjab University and Masters in Mechanical Engineering from University of Roorkee in 1984. He received Ph.D. degree in 1991 in Mechanical Engineering from University of Roorkee, India. He has about 25 years research/teaching and industrial experience and more than 150 publications in international/national journals/conferences. He has guided number of students for their undergraduate projects, master's dissertations and Ph.D. degrees. His fields of interests are system behavior in industry, maintenance engineering, reliability analysis and industrial engineering.

Journal of Alloys and Compounds 645 (2015) 171–177

Contents lists available at ScienceDirect

Journal of Alloys and Compounds

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

Consequence of pH variation on the dielectric properties of Cr-doped lithium ferrite nanoparticles synthesized by the sol–gel method

Manish Srivastava [a,b], R.K. Mishra [c], Jay Singh [d], Neha Srivastava [e], Nam Hoon Kim [a,*], Joong Hee Lee [a,*]

[a] Advanced Materials Research Institute for BIN Convergence Technology (BK Plus Global Program), Department of BIN Convergence Technology, Chonbuk National University, Jeonju, Jeonbuk 561-756, Republic of Korea
[b] Department of Physics & Astrophysics, University of Delhi, Delhi 110007, India
[c] Department of Physics, Motilal Nehru National Institute of Technology, Allahabad, Allahabad 211004, India
[d] Department of Applied Chemistry & Polymer Technology, Delhi Technological University, Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India
[e] Department of Chemical Engineering and Technology, Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh 221005, India

ARTICLE INFO

Article history:
Received 7 October 2014
Received in revised form 28 April 2015
Accepted 4 May 2015
Available online 9 May 2015

Keywords:
Lithium ferrite
Nanoparticles
Dielectric properties
Sol–gel

ABSTRACT

Nanocrystalline Cr-doped lithium ferrites have been synthesized at different pH values through sol–gel process. Dielectric properties of the synthesized samples were studied via impedance measurements in the frequency range of 100 Hz–7 MHz at room temperature. It was found that the dielectric properties strongly depended on the pH value, and the sample synthesized at pH 3.5 exhibited the highest dielectric constant, on the order of $\sim 10^7$. In contrast, the sample synthesized at pH 11.5 exhibited the lowest dielectric constant. The occurrence of a peak in the dielectric loss spectra was found to be pH dependent, whereas the AC conductivity increased with frequency and was the highest in the sample synthesized at pH 3.5. The impedance spectroscopy and modulus formulism were also employed to investigate the charge transfer and relaxation dynamic properties. Consequently, dielectric properties of the samples were easily tuned by varying the pH during the synthesis process.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Among the different types of spinel ferrites, lithium (Li)-based ferrites have particular importance and have received considerable attention from researchers. Li-ferrites possess unique properties such as high Curie temperature, square hysteresis loops, high saturation magnetization, as well as they are excellent replacement for garnet due to their low cost and superior performance, which has prompted further study of Li and its substituted ferrites [1]. The above mentioned properties make them suitable candidates for microwave [2] and electromagnetic components like phase shifters, circulators, isolators, deflection yokes, transformer cores, and memory devices [3]. Cubic spinel-type Li-ferrites can be represented by the general formula AB_2O_4 [4], where A and B are the metal cations, which exist in the tetrahedral and octahedral voids whereas oxygen anions exist in the cubic close-packed structure. Moreover, in these ferrites materials cation distribution over the tetrahedral and octahedral voids is mainly responsible for the unusual physical properties, which depend on the synthesis procedure and several parameters [5,6]. Among various parameters, pH is an important factor that influences the distribution of metal cations in the tetrahedral and octahedral voids, and moreover, growth of grains and their boundaries are also affected which causes different physical properties in the final products [7,8]. It has been reported that variation of pH during the synthesis process can influenced the thermodynamics of chemical reactions which results different physical properties of the ferrites materials [7] Shirsath et al. synthesized $CoDy_{0.1}Fe_{1.9}O_4$ nanoparticles at four different pH values in the range of 2.5–10 [8]. These authors have reported that pH value play a significant role to synthesize single phase $CoDy_{0.1}Fe_{1.9}O_4$ nanoparticles. Besides, variation in pH during the synthesis process influenced the structural phase, size, structural morphology, bulk density and cation distribution of nanoparticles which resulted due to the variation of combustion rate of the chemical solution during the synthesis. Numerous studies on Li and their doped ferrites have been reported [9–17]. Mane et al. synthesized chromium doped lithium ferrite nanoparticles by sol–gel auto combustion and thoroughly studied the cation distribution, structural and magnetic properties of the samples [18]. However, to the best of our knowledge, a study describing the dielectric properties of chromium substituted Li-ferrite, synthesized at different pH values has yet to be reported. Previously, we studied the effect of pH on the magnetic properties of the

* Corresponding authors. Tel.: +82 63 270 2342; fax: +82 63 2702341.
E-mail addresses: nhk@chonbuk.ac.kr (N.H. Kim), jhl@chonbuk.ac.kr (J.H. Lee).

http://dx.doi.org/10.1016/j.jallcom.2015.05.017
0925-8388/© 2015 Elsevier B.V. All rights reserved.

Cr-substituted Li-ferrites nanoparticles and observed that the pH value remarkably influenced magnetic properties of the final product. A relation between the cation distribution and magnetic properties was investigated through combined studies of Mossbauer spectroscopy and Vibrating Sample Magnetometer (VSM) measurements [19]. Therefore, as an extension to our previous work, the aim of the present work is to study the influence of pH on dielectric properties of Cr-substituted Li-ferrites nanoparticles. The variations of dielectric constant, loss tangent, AC conductivity, impedance and modulus have been studied as a function of frequency. Detailed knowledge of the effect of pH on the dielectric properties might be useful for producing materials with tunable dielectric properties, which is beneficial in practical applications [19,20].

2. Experimental

2.1. Materials and synthesis method

Li acetate (extra pure, Merck), ferric nitrate (AR, Qualigens), chromium nitrate, 25% aqueous solution of ammonia (Anala R, BDH), and 70% aqueous solution of glycolic acid were used as raw materials. Li ferrite nanoparticles with a compositional formula of $Li_{0.5}Cr_{0.1}Fe_{2.4}O_4$ were synthesized by following the sol–gel method. The detail synthesis procedure is reported in an earlier study [19]. Four samples at different pH values of 3.5, 6.5, 9.5, and 13.5 were synthesized. A schematic representation of the synthesis process is shown in Fig. 1(a).

2.2. Characterization

The details of the structural analysis, Mossbauer study, and magnetic properties of the synthesized sample are reported elsewhere [19]. Surface morphologies of the synthesized samples are investigated through scanning electron microscope (SEM) using FEI Quanta-200 MK2. Dielectric properties of the samples were measured at room temperature in the frequency range of 100 Hz to 7 MHz. AC measurements were carried out using a precision impedance analyzer (Agilent Technologies, Model: 4294A). Measurements were performed by making pellets from a fine powder by introducing a hydraulic pressure of ~15 MPa. The diameter and thickness of the pellets were 12 mm and 3 mm, respectively, and they were coated with a thin layer of silver paste for making a good electrical contact. Fig. 1(b) shows the experimental arrangement for the dielectric measurement.

3. Results and discussion

3.1. Dielectric constant

In general, the dielectric constant of a material is related to the phenomena of polarization and it is expected to be high where the polarization is high. Frequency dependent variations of real and imaginary part of the dielectric constant in the range of 100 Hz–7 MHz measured at room temperature are shown in Fig. 2(a) and (b). The relative, real and imaginary parts of the dielectric constants are represented by the following equations [21].

$$\varepsilon^* = \varepsilon' + i\varepsilon'' \quad (1)$$

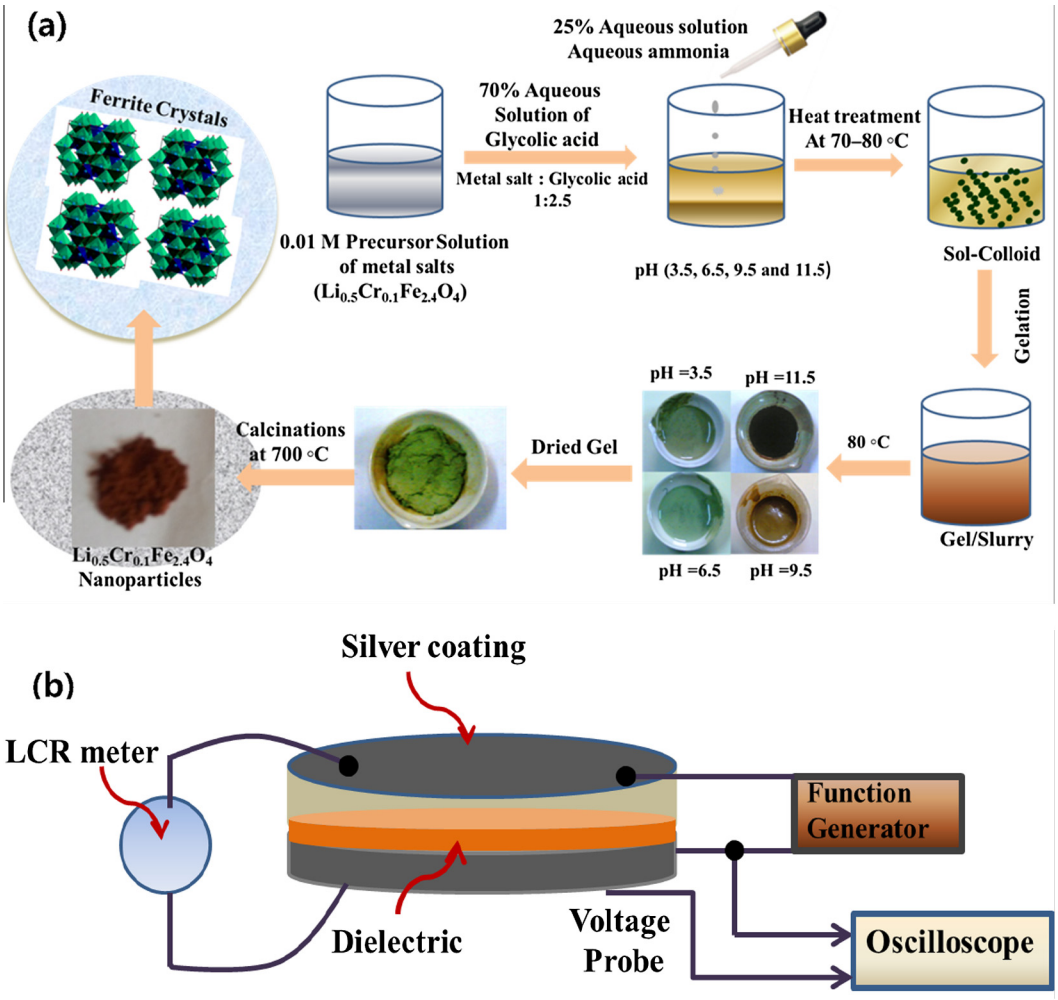


Fig. 1. Schematic illustration of sample synthesis (a) and experimental arrangement for the measurements of the dielectric properties (b).

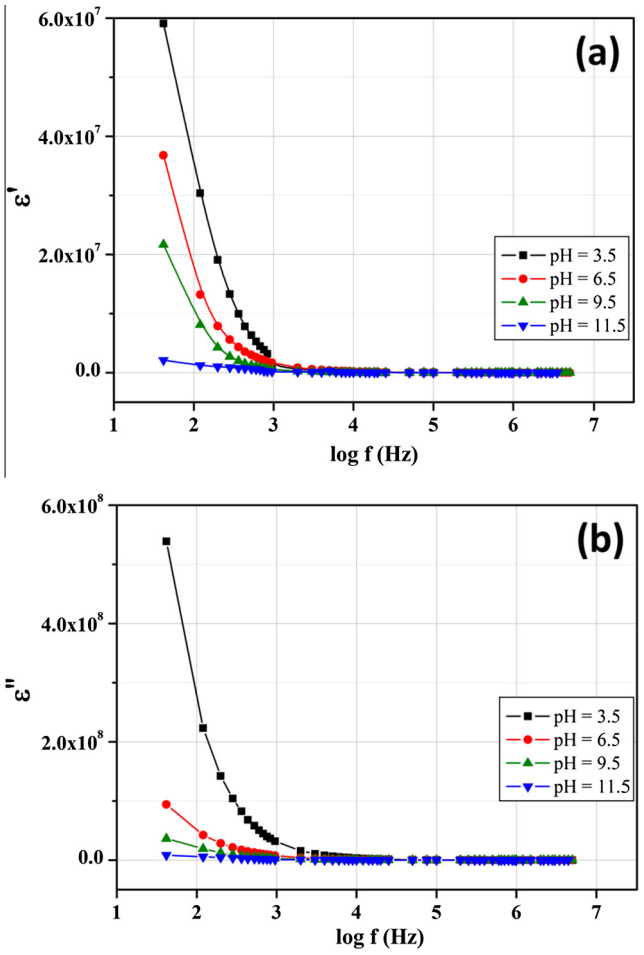


Fig. 2. Frequency dependence of ε' (a) and ε'' (b) for samples synthesized at different pH values.

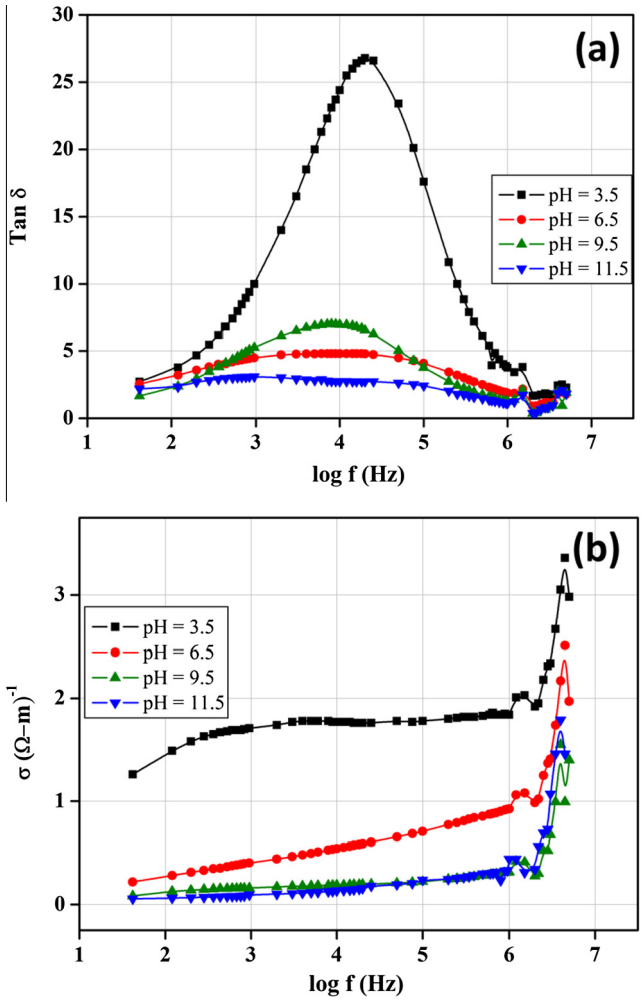


Fig. 3. Frequency dependence of the tan δ (a) and AC electrical conductivity (b) for the samples synthesized at different pH values.

$$\text{where} \quad \varepsilon' = \frac{Z''}{\omega C_o (Z'^2 + Z''^2)} \tag{2}$$

$$\varepsilon'' = \frac{Z'}{\omega C_o (Z'^2 + Z''^2)} \tag{3}$$

Here Z' (=$Z\cos\phi$), Z'' (=$Z\sin\phi$), ϕ, ω, and C_o (=ε_o (A/t)) are the real, imaginary parts of the complex impedance Z, phase angle, angular frequency, and vacuum capacitance of the cell, respectively.

In the low frequency region, dielectric dispersion can be observed for all the samples where both real and imaginary dielectric constant decreases rapidly with increasing frequency and exhibits almost frequency independent behavior in the high frequency region. The independence of the dielectric constant on frequency in the high frequency region (3 MHz–7 MHz) is due to the influence of a large relaxation time; on the other hand, at low frequencies dielectric constants are high. This may be due to the space charge polarization at the grain boundaries, which generate a potential barrier; therefore, accumulation of charge at the grain boundary can occur, which leads to higher values of the dielectric constant [22]. On the other hand, the dielectric behavior of the real part depends on the rapid polarization in the samples. This behavior may be due to the existence of a large number of trap states between the valence band, and conduction band which arises due to the formations of Frenkel or Schottky defects in the sample during crystal growth. These trap states are mainly responsible for the existence of different charges having different time constants. The imaginary part of the dielectric constant depends on the varying electric field. If the electric polarizations are independent of the varying electric field, loss in the charge storing capacity of materials occurs. Moreover, it can also be seen that the drop in the imaginary part of the dielectric constant is more pronounced than the real part (Fig. 2(b)). Another explanation of the variation in the dielectric constant may be understood on the basis of Maxwell–Wagner theory and Koop's model [23–26]. It is suggested that , in the case of ferrites, a high value of the dielectric constant may arise due to the space charge polarization near the grain boundary, since the dielectric structure of ferrites have well conducting grains separated by highly resistive grain boundaries, where the drop in applied voltage causes a build-up of space charge polarization. This phenomenon can produce a large value of dielectric constant. Furthermore, the polarization mechanism involves the exchange of electrons between the ions of the same element where the different dipoles have different relaxation times. During the exchange process, electrons pass through the grain and its boundary, which causes the accumulation of electrons near the grain boundary and produces space charge polarization. A rapid decrease in dielectric constant at low frequency and an almost constant value in the high frequency region are attributed to the grain boundary effect. In the high frequency region, the grains play a crucial role, and the hopping of electrons also cannot follow the high frequency AC field; as a result, dielectric constant decreases and

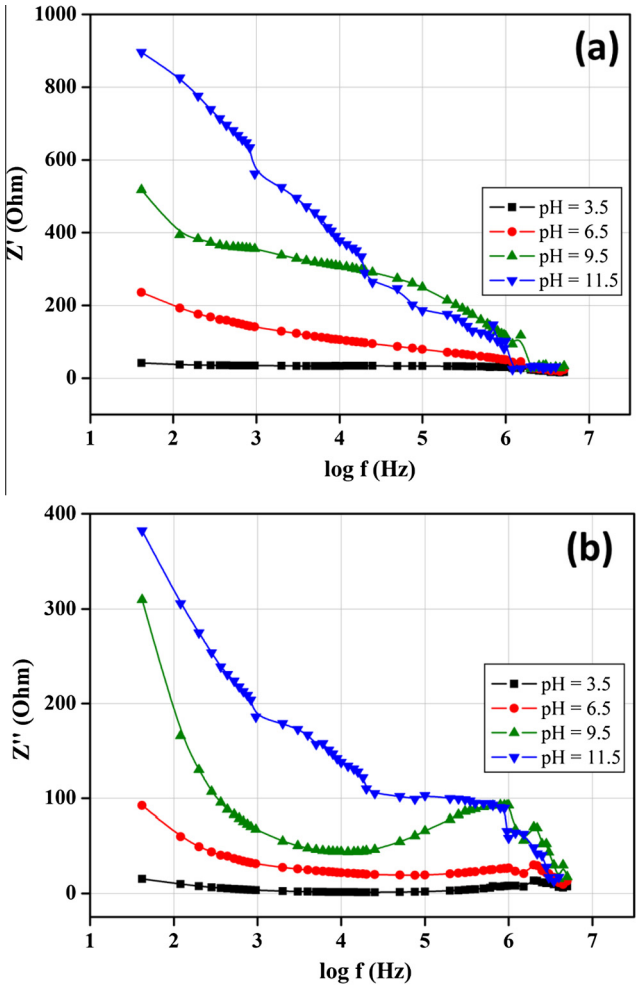


Fig. 4. Frequency dependence of the Z' (a) and Z'' (b) for samples synthesized at different pH values.

becomes frequency independent [27]. In the present study, the real and imaginary parts of the dielectric constant decrease with increasing pH value. Combining this result with the results of our earlier study, we conclude that an increase in pH value during the synthesis process led to the distribution of more Fe^{+3} ions on octahedral sites than on tetrahedral sites (Table S1), which can possibly lower the polarizability, which, in turn, decreased the dielectric constant.

3.2. Tangent loss

The tangent loss (Eq. (4)) is the measurement of the energy loss within the dielectric materials, arising when the polarization lags behind the applied AC field or is caused by domain wall resonance, due to crystal impurities and defects [28]. The variations of $\tan\delta$ with frequency for the samples synthesized at different pH values are shown in Fig. 3(a). It can be seen that dielectric loss shows a peak for all the samples along with a relative shift in the maxima. Further, the height of the peak differs for samples synthesized at different pH values. The presence of a peak in the $\tan\delta$ versus frequency curve is correlated with the conduction mechanism and the dielectric behavior in the spinel ferrites. The peak is observed when the hopping frequency of the electrons is in resonance with the frequency of applied electric field, i.e., the maximum electric energy is transferred to the electrons and the loss increases dramatically at resonance conditions. Similar kinds of observations have also been reported in an earlier study [22].

$$\tan\delta = \varepsilon''/\varepsilon' \tag{4}$$

The condition for maximum dielectric loss of a dielectric material is given by the relation

$$\omega\tau = 1 \tag{5}$$

where $\omega = 2\pi f_{max}$, and τ is the relaxation time, which is related to the jumping probability (p) per unit time, $\tau = 1/2p$.

Thus, the shift in the maxima towards the low frequency region indicates a decrease in jumping probability for sample synthesized at higher pH values, which influences the distribution of metal cations in the tetrahedral and octahedral voids and causes variation in morphological growth and particle size as well (Fig. S1) [29–31]. Moreover, this shift may also be associated with the distributions of Fe^{3+} ions at B-sites, which influence the probability of jumping and are consequently responsible for the polarization. It can also be seen that the sample synthesized at pH 3.5 exhibits the highest dielectric loss which is attributed to the cation distribution at tetrahedral and octahedral sites. As presented in (Table S1), it is clear that distribution of Fe^{3+} ions at B-sites were least favorable for the sample synthesized at pH 3.5 whereas, most favorable for the sample synthesized at pH, 11.5. On the other word, highest number of Fe^{3+} ions at A-sites (sample synthesized at pH 3.5) may easily contribute towards the hopping of electron at resonance conditions, resulted in greater dielectric loss compared to the sample synthesized at higher pH value.

3.3. Conductivity

In order to understand the conduction mechanism, variations in AC conductivity as the function of frequency have been investigated and graphs are shown in Fig. 3(b). In ferrites, electrical conductivity is mainly followed by hopping of electrons between the ions of the same elements, which are randomly distributed over the tetrahedral and octahedral sites. Usually, AC conductivity is directly proportional to the frequency in disordered solids. Austin et al. have discussed the electrical conduction mechanism in terms of polarons and an electron hopping model [32]. The AC conductivity decreases with frequency in large polaron hopping, whereas it increases for small polaron hopping. In general, the total conductivity can be given as follows [33]:

$$\sigma(\omega,T) = \sigma_{dc}(T) + \sigma_{ac}(\omega,T) = \sigma_{dc}(T) + B(T)\omega^{n(T)} \tag{6}$$

where σ_{dc} is the DC conductivity, σ_{ac} is the ac conductivity, and $B(T)$ is a temperature dependent quantity. The AC conductivity obeys the frequency dependence universal power law given as

$$\sigma_{ac}(\omega,T) = B(T)\omega^{n}(T) \tag{7}$$

where both B and n are the constant and depend upon temperature and composition. B has the units of electrical conductivity, whereas n is dimensionless. The value of n varies from 0 to 1. Additionally, when $n \leqslant 1$, the conductivity is frequency dependent (AC conductivity), however, for $n = 0$ it is DC. In this work, we measured the frequency dependent AC conductivity. It can be seen from the graph (Fig. 3(b)) that the electrical conductivity gradually increases with increasing frequency, which is in agreement with Eq. (7) and agrees generally with the behavior of semiconducting materials [21]. This phenomenon can be attributed to two possible mechanisms: (1) the electrical energy associated with the applied AC field helps in the liberation of trapped charges and their transfer between the different localized states, and (2) enhanced dielectric relaxation of the polarization of Li-ferrite nanoparticles occurs in the higher frequency range. These observations are also compatible with earlier studies, where it was suggested that the increase in AC conductivity

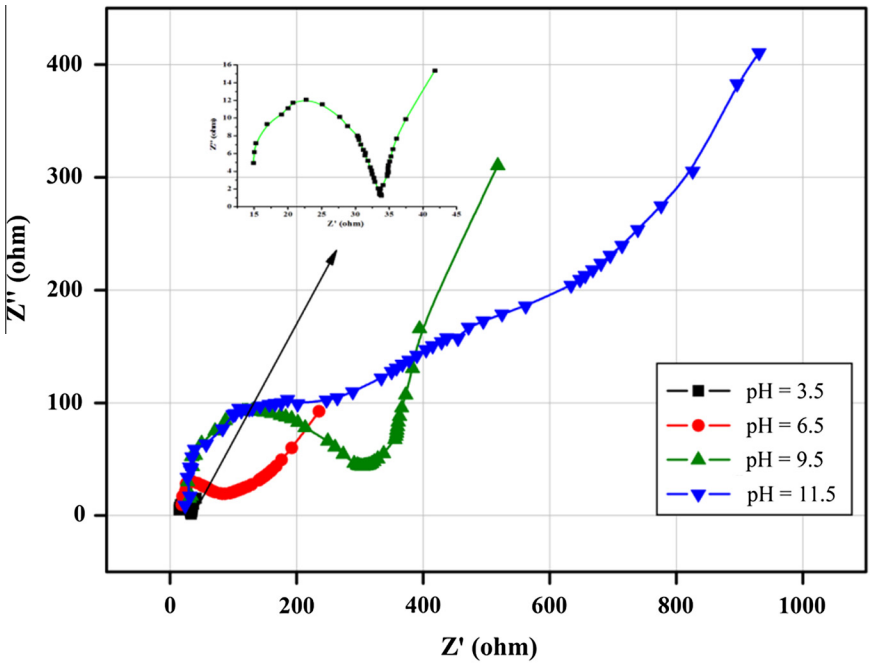


Fig. 5. Nyquist plot for the samples synthesized at different pH values (the inset shows the magnified portion of the Nyquist plot for the sample synthesized at pH 3.5).

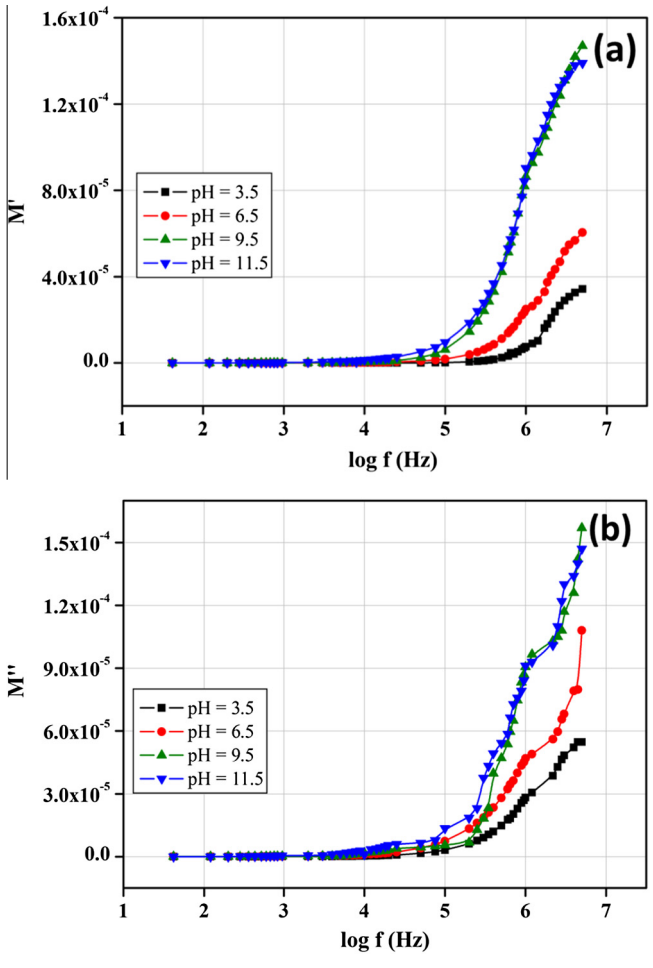


Fig. 6. Frequency dependence of M' (a) and M'' (b) for samples synthesized at different pH values.

is the consequence of the series resistance effect [28,34,35]. It can also be seen that the value of conductivity has a remarkable influence on pH variation during the synthesis process and can be correlated with the dielectric measurement (combining Eqs. (2), (3) and (7)), where the dielectric constant is observed to be maximum for the sample synthesized at pH 3.5 (lowest number of Fe^{3+} ions at B-sites) and minimum for the sample synthesized at pH 11.5 (highest number of Fe^{3+} ions at B-sites) (Table S1). The value of 'n' was calculated in the range of 0.29–0.99, according to literature [36], which suggests the AC conduction in the samples is due to the hopping, of charges.

3.4. Impedance spectroscopy

To investigate the transport properties, impedance measurements have been carried out in the frequency range of 100 Hz–7 MHz at room temperature. The complex impedance of the material is given as [33]:

$$Z^* = Z'(\omega) - iZ''(\omega) \quad (8)$$

where Z' and Z'' are the real and imaginary parts, respectively.

Fig. 4(a) shows the variation of the real part of the impedance (Z') as a function of frequency for the samples synthesized at different pH values. In the low frequency range the value of Z' is high, whereas a monotonous decrease in the value of Z' can be seen with increasing frequency. Larger values of Z' at low frequency suggest larger polarization effects. On the other hand, the decreasing trend in Z' with increasing frequency results in an increased AC conductivity at higher frequency. Additionally, in the high frequency range, the merging of Z' is possibly due to the release of space charge polarization which reduces the barrier properties in the materials [37]. Fig. 4(b) shows the variation of Z'' with frequency. It can be seen that the value of Z'' first decreases up to a certain frequency, after which it starts increasing until reaching a peak in the high frequency range. These phenomena can be correlated with the frequency dependent relaxation behavior in the samples. Furthermore, values of Z' and Z'' in the low frequency range for all the samples were found to increase with increasing pH, in

contrast to the dielectric value where it was observed to be decreasing with increasing pH value. These observations further suggest that the impedance measurements are consistent with the dielectric results.

3.5. Nyquist plot

The analysis of complex impedance can provide information about electrical and structural properties of the materials. Usually, the impedance spectrum is represented as a Z' versus Z'' plot, referred to as a Nyquist plot. Depending upon the nature of the materials, this plot can give two semicircles, where the semicircle at low frequency corresponds to the resistance of the grain boundary and the other at high frequency corresponds to the bulk or grain properties. The Nyquist plots of the samples synthesized at different pH values are shown in Fig. 5. It can be seen that all samples exhibit a semicircle, and the intercept of the arc on the real axis gives an estimation of the grain resistance. Samples synthesized at higher pH values display larger arcs, and therefore have a relatively higher resistance. On the other hand, the incomplete semicircle also suggests that the grain boundary resistance is out of measurement range [38]. These observations are further supported by dielectric and AC conductivity results.

3.6. Modulus spectroscopy

Modulus spectroscopy is an important and alternate formalism for analyzing the electrical response. It can provide more information about the mechanism of electrical transport, ion dynamics and relaxation in conductivity. The components of the complex electrical modulus can be represented as [39]:

$$M' = 2\pi f C_o Z'' \tag{9}$$

$$M'' = 2\pi f C_o Z' \tag{10}$$

Fig. 6(a) and (b) shows variation in the real M' and imaginary M'' part of the modulus as a function of frequency, respectively. It can be seen that, in the low frequency region, both M' and M'' are very small (approaching zero). However, M' and M'' show an increase with frequency, where both exhibit a tendency to saturate at a maximum asymptotic value designated as M_∞. This phenomenon may be due to the lack of restoring force, which governs the mobility of charges under the action of an induced electric field. It can also be seen that ε' and ε'' decrease with increasing frequency, whereas M' and M'' increase. It is suggested that ε' approaches infinity and M' also approaches infinity, corresponding to $M_\infty = 1/\varepsilon_\infty$. M' and M'' approach zero at low frequency due to the removal of electrical polarization [21]. It is clear that at high frequency interface state cannot follow the AC signal, therefore M' and M'' achieve a maximum value.

4. Conclusions

Cr-doped Li-ferrite nanoparticles were synthesized by sol–gel method at four different pH values. Dielectric properties of samples were significantly influenced by the different pH values, and the samples synthesized at lower pH exhibited a relatively higher value of dielectric constant. All samples exhibited a reasonably high dielectric constant, on the order of $\sim 10^7$. The resonance peak was observed in the frequency dependence of $\tan\delta$ plot and the position of the peak shifts towards the lower frequency when samples were synthesized at higher pH values. AC conductivity increased with frequency, and it was found to be highest in the sample synthesized at pH 3.5. In addition, impedance spectroscopy and modulus formulism were also employed in order to investigate the charge transfer and relaxation dynamic properties in the samples. Finally, dielectric properties were simply tuned by varying the pH value during the synthesis process.

Acknowledgements

This study was supported by the Converging Research Center Program (2014M3C1A8048834) and the Basic Research Laboratory Program (2014R1A4A1008140) through the Ministry of Science, ICT & Future Planning of Korea. MS acknowledges the DST, Government of India for a DST-INSPIRE Fellowship [IFA13-MS-02] 2014. JS acknowledges the DST, Government of India for a DST-INSPIRE Fellowship [IFA-13 CH-105] in 2013 and a DST Young Scientist award (CS-393/2012). Author N.S. thankfully acknowledges to DST, New Delhi, India for providing the Women scientist-B fellowship (SEED/DISHA/WOSB/047/2012/G) and department of chemical engineering and technology, IIT (BHU), Varanasi, India.

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.jallcom.2015.05.017.

References

[1] P.V.B. Reddy, B. Ramesh, C.G. Reddy, Phys. B 405 (2010) 1852–1856.
[2] X. Cao, K. Sun, C. Sun, L. Leng, J. Mag. Mag. Mat. 321 (2009) 2896–2901.
[3] E.D. Fazio, P.G. Bercoff, S.E. Jacobo, J. Mag. Mag. Mat. 323 (2011) 2813–2817.
[4] J. Slama, M. Sokan, A. Gruskova, R. Dosoudil, V. Jancarik, J. Degmov, J. Mag. Mag. Mat. 326 (2013) 251–256.
[5] B. Ramesh, M.L. Rao, J. Alloys Comp. 551 (2013) 527–530.
[6] M.F.A. Hilli, S. Li, K.S. Kassim, Mater. Sci. Eng. B 158 (2009) 1–6.
[7] (a) L. Aditya, J. Nanda, I. Samajdar, N. Venkataramani, S. Prasad, J. Mag. Mag. Mat. 321 (2009) 3373–3379;
(b) A. Shanmugavani, R. Kalai Selvan, Samar Layek, Leonid Vasylechko, C. Sanjeeviraja, Mater. Mater. Res. Bull., http://dx.doi.org/10.1016/j.materresbull.2015.04.008.
[8] M. Srivastava, J. Singh, M. Yashpal, D.K. Gupta, R.K. Mishra, S. Tripathi, S. Tripathi, Animesh K. Ojha, Carbohydr. Polym. 89 (2012) 821–829.
[9] S.E. Shirsath, M.L. Mane, Y. Yasukawa, X. Liu, A. Morisako, J. Nanopart. Res. 15 (2013) 1976–1988.
[10] M. Srivastava, A.K. Ojha, S. Chaubey, P.K. Sharma, A.C. Pandey, Mater. Sci. Eng. B 175 (2010) 14–21.
[11] S.E. Shirsath, R.H. Kadam, M.L. Mane, A. Ghasemi, Y. Yasukawa, X. Liu, A. Morisako, J. Alloys Comp. 575 (2013) 145–151.
[12] N.I. Abu-Elsaad, J. Mol. Struct. 1075 (2014) 546–550.
[13] M.A. Dar, V. Verma, V. Pandey, R. Sonkar, W.A. Siddiqui, R.K. Kotnala, Integr. Ferroelectr. 119 (2010) 135–142.
[14] R.G. Kharabe, B.K. Bammannavar, B.K. Chougule, J. Nanosci. Nanotechnol. 11 (2011) 2735–2738.
[15] V. Mohanty, R. Cheruku, L. Vijayan, G. Govindaraj, J. Mater. Sci. Technol. 30 (2014) 335–341.
[16] P.R. Arjunwadkar, R.R. Patil, J. Alloys Comp. 611 (2014) 273–277.
[17] S.A. Mazen, N.I. Abu-Elsaad, J. Mag. Mag. Mat. 324 (2012) 3366–3373.
[18] D.R. Mane, S. Patil, D.D. Birajdar, A.B. Kadam, S.E. Shirsath, R.H. Kadam, Mater. Chem. Phys. 126 (2011) 755–760.
[19] M. Srivastava, S. Layek, J. Singh, A.K. Das, H.C. Verma, A.K. Ojha, N.H. Kim, J.H. Lee, J. Alloys Comp. 591 (2014) 174–180.
[20] S.K. Gurav, S.E. Shirsath, R.H. Kadam, D.R. Mane, Powder Technol. 235 (2013) 485–492.
[21] H.M. Chenari, A. Hassanzadeh, M.M. Golzan, H. Sedghi, M. Talebian, Curr. Appl. Phys. 11 (2011) 409–413.
[22] M.A. Dar, K.M. Batoo, V. Verma, W.A. Siddiqui, R.K. Kotnala, J. Alloys Comp. 493 (2010) 553–560.
[23] C.G. Koops, Phys. Rev. 83 (1951) 121–124.
[24] J.C. Maxwell, Electricity and Magnetism, Oxford University Press, London, 1973.
[25] K.W. Wagner, Ann. Phys. (Leipzig) 40 (1913) 53.
[26] K.W. Wagner, Arch Elektrotech 2 (1914) 371–387.
[27] M.L. Mane, S.E. Shirsath, V.N. Dhage, K.M. Jadhav, Nucl. Instr. Meth. Phys. Res. B 269 (2011) 2026–2203.
[28] N. Singh, A. Agarwal, S. Sanghi, Curr. Appl. Phys. 11 (2011) 783–789.
[29] M. Maisnam, S. Phanjoubam, Solid State Commun. 152 (2012) 320–323.
[30] M. Hashim, S.E. Shirsath, R.K. Kotnala, S.S. Meena, S. Kumar, A. Roy, R.B. Jotania, P. Bhatt, R. Kumar, J. Alloys Comp. 573 (2013) 198–204.

[31] D. Rathore, R. Kurchania, R.K. Pandey, Int. J. Min. Met. Mater. 21 (2014) 408–414.
[32] I.G. Austin, N.F. Mott, Adv. Phys. 18 (1969) 41–102.
[33] A. Azam, A.S. Ahmed, M.S. Ansari, M. Shafeeq, A.H. Naqvi, J. Alloys Comp. 506 (2010) 237–242.
[34] A.R. Babar, S.S. Shinde, A.V. Moholkar, K.Y. Rajpure, J. Alloys Comp. 505 (2010) 743–749.
[35] M. Srivastava, J. Singh, R.K. Mishra, M.K. Singh, A.K. Ojha, M. Yashpal, S. Sudhanshu, Curr. Appl. Phys. 14 (2014) 980–990.
[36] M.S. Samuel, J. Koshy, A. Chandran, K.C. George, Curr. Appl. Phys. 11 (2011) 1094–1099.
[37] A. Shukla, R.N.P. Choudhary, Curr. Appl. Phys. 11 (2011) 414–422.
[38] P.P. Hankare, R.P. Patil, U.B. Sankpal, K.M. Garadkar, R. Sasikala, A.K. Tripathi, I.S. Mulla, J. Mag. Mag. Mat. 322 (2010) 2629–2633.
[39] B. Unal, Z. Durmus, A. Baykal, M.S. Toprak, H. Sozeri, A. Bozkurt, J. Alloys Comp. 509 (2011) 8199–8206.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

Adiabatic Vedic Multipler Design Using Chinese Abacus Approach

Diksha Ruhela[1], Dr. Malti Bansal[2]

M.Tech (VLSI), Electronics & Communication Engineering, Delhi Technological University, Delhi, India[1]

Assistant Professor, Electronics & Communication Engineering, Delhi Technological University, Delhi, India[2]

Abstract: In this paper, we discuss the implementation of Vedic multiplier with Chinese Abacus adder design, using Reversible Logic Gates. The power consumption of the Vedic multiplier is low as it generates all partial products and their sum in one step. Since high radix of Chinese Abacus adder reduces the carry propagation delay, it is observed here that the proposed design increases the speed of operation manifold. The proposed work is implemented on the Xilinx FPGA device, Spartan-3E. The results show that multiplier implemented using Chinese Abacus approach is quite efficient in terms of area, time and speed

Keywords: Vedic Multiplier, Chinese Abacus Adder, propagation delay, Reversible Logic.

I. INTRODUCTION

Multiplier is one of the important blocks in almost all the arithmetic logic units. These multipliers are mostly used in the fields of Digital Signal Processing (DSP), Fast Fourier Transforms, convolution, filtering and microprocessor applications. A system's performance is generally determined by the performance of the multiplier, because the multiplier is generally the slowest element in the system. Furthermore, it is generally the most area consuming. Hence, optimizing the speed and area of the multiplier is a major design issue. A high speed and area efficient multiplier can be achieved by using Vedic mathematics. In this work, we have implemented the Vedic multiplier using Chinese Abacus Adder with and without using Reversible logic gates.

Reversible logic is one of the promising fields for future low power design technologies. Since one of the requirements of all DSP processors and other embedded devices is to minimize power dissipation, multipliers with high speed and lower dissipation are critical. This paper proposes an implementation of Reversible *Urdhva Tiryakbhayam* (UT) Multiplier which consists of two cardinal features. One is the fast multiplication feature derived from Vedic algorithm *Urdhva Tiryakbhayam* and another is the reduced heat dissipation by the virtue of implementing the circuit using reversible logic gates. The paper is partitioned into six sections. Section II gives introduction on reversible logic. Section III explains the *Urdhva Tiryakbhayam* (UT) algorithm. Section IV explains the Chinese Abacus Adder algorithm, Section V elaborates on the design aspects of Reversible UT Multiplier. Section VI gives the simulation results. Conc1usions and references follow

II. REVERSIBLE LOGIC

Reversible logic is a promising computing design paradigm which presents a method for constructing computers that produce no heat dissipation. Reversible computing emerged as a result of the application of quantum mechanics principles towards the development of a universal computing machine. Specifically, the fundamentals of reversible computing are based on the relationship between entropy, heat transfer between molecules in a system, the probability of a quantum particle occupying a particular state at any given time, and the quantum electrodynamics between electrons when they are in proximity. The basic principle of reversible computing is that an objective device with an identical number of input and output lines will produce a computing environment where the electrodynamics of the system allow for prediction of all future states based on known past states, and the system reaches every possible state, resulting in no heat dissipation. A reversible logic gate is an N-input, N-output logic device that provides one to one mapping between the input and the output. It not only helps us to determine the outputs from the inputs but also helps us to uniquely recover the inputs from the outputs.

The following are the important design constraints for reversible logic circuits:

1. Reversible logic gates do not allow fan-outs.
2. Reversible logic circuits should have minimum quantum cost.
3. The design can be optimized so as to produce minimum number of garbage outputs.
4. The reversible logic circuits must use minimum number of constant inputs.
5. The reversible logic circuits must use a minimum logic depth or gate levels.

The basic reversible logic gates encountered during the design are listed below:

1. Feynman Gate [1]:

It is a 2x2 gate and its logic circuit is as shown in the figure. It is also known as Controlled Not (CNOT) Gate. It has quantum cost one and is generally used for Fan Out purposes.

2. Peres Gate [2]:

It is a 3x3 gate and its logic circuit is as shown in the figure. It has quantum cost four. It is used to realize various Boolean functions such as AND, XOR.

Copyright to IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

3. Fredkin Gate [3]:

It is a 3x3 gate and its logic circuit is as shown in the figure. It has quantum cost five. It can be used to implement a Multiplexer.

4. HNG Gate[4]:

It is a 4x4 gate and its logic circuit is as shown in the figure. It has quantum cost six. It is used for designing ripple carry adders. It can produce both sum and carry in a single gate thus
minimizing the garbage and gate counts.

5. NFT Gate[5]:

The 3*3 New Fault Tolerant gate (NFT) with quantum cost of 5 has worst case delay of 3 and it has better correction capability. The output states map to the inputs in this manner.

A = A XOR B,
B = AC' XOR B'C, and
C = AC' XORB C

6. F2G Gate[5]

The 3*3 Feynman double gate gate with quantum cost of 2 has worst case delay of 3 and it has better correction capability

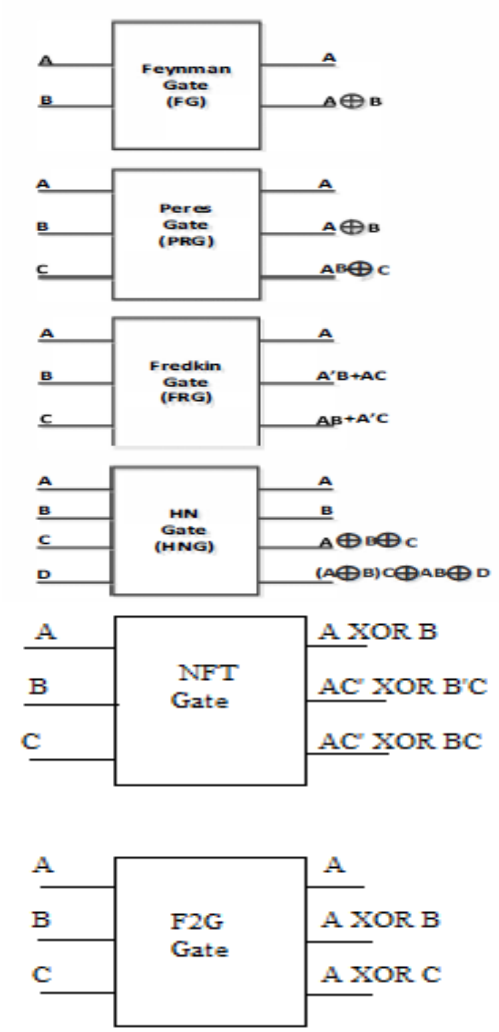


Figure1: Reversible Logic Gates[1][2][3][4][5]

III. URDHAVA – TIRYAGBHYAM

The multiplier is based on an algorithm *Urdhva Tiryakbhyam* (Vertical & Crosswise) of ancient Indian Vedic Mathematics. *Urdhva Tiryakbhyam* Sutra is a general multiplication formula applicable to all cases of multiplication. It literally means "Vertically and crosswise". It is based on a novel concept through which the generation of all partial products can be done with the concurrent addition of these partial products.

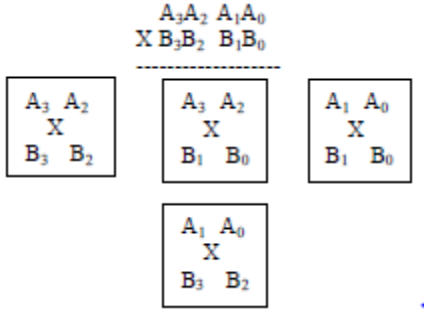


Figure 2: Block diagram of 4x4 vedic multiplication.

For NXN multiplication unit, we require four N/2 bit multipliers and three N bit full adders. 4x4 bit multiplication has been shown in Fig.3. High speed of multiplier depends mainly upon speed of the adder units used.

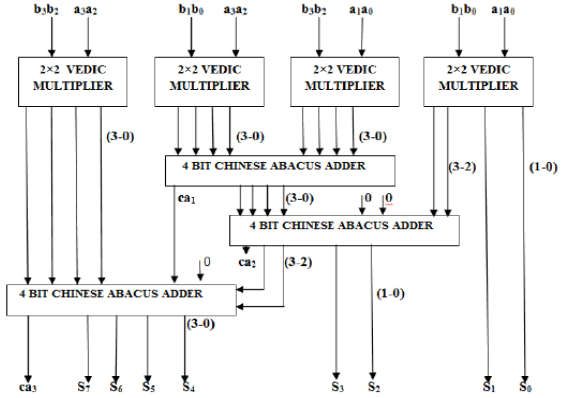


Figure3: Block diagram of proposed 4x4 Vedic multiplier using Chinese Abacus Adder

IV. CHINESE ABACUS

Chinese Abacus is the ancient approach used for performing mathematical calculation in many parts of the world, especially in China. It is still used by many small enterprises for their day to day work. Due to its high speed of operation, it give high competition to electronic pocket calculator. Similar efficiency can be attained in electronic version.

The first digital arithmetic circuit employing the Chinese abacus approach were proposed by Maloberti et al. [6-8]. The architecture of radix-4 abacus adder is presented in [6]. The architecture of radix-2 abacus adder is presented in [7] and [8]. The most severe drawback of Maloberti's adder in radix-4 design is the serial addition of each bead by the Shift-Up module. The serial addition of each bead by the Shift-Up module can be improved by the parallel addition

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

module of the proposed architecture. The serial addition of each bead by the Shift-Up module will moreover induce more than nine pass transistor chains, due to its counting method. The delay time of the signal waveform becomes long.

Radix-4 adder helps in reducing the excessive delay by reducing the number of carries [9]. A basic column element of the Chinese abacus is depicted in figure 4(a). The configuration showed in figure 4(a) represents the number six. Each column element has one higher bead with a weight, of five and four lower beads with a weight of one. The key feature of the Chinese abacus is the use of one bead with weight five. This allows the operator to minimize the transmission of rests.

In this paper, we use abacus adder in each column having three beads higher and three beads lower, each having weight of four and one each, respectively. The basic element is able to represent decimal number in the range 0 to 15. Figure 4(b) depicts the abacus representation of decimal number 9. In this methodology, the 4-bit adder only contributes two ripple carries. One is internal carry from lower beads to higher beads, the other is external carry from this column element to a next neighbour column element.

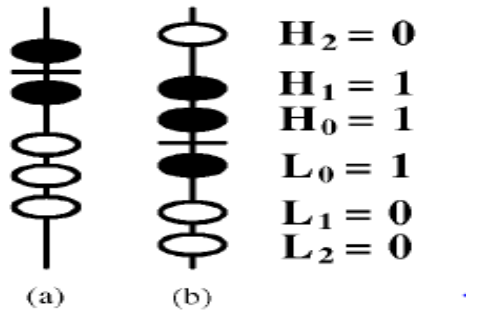


Figure 4. (a) Chinese abacus coding with base 10 of the decimal number 6, (b) the proposed Chinese abacus adder coding with base 16 of the decimal number 9 [7].

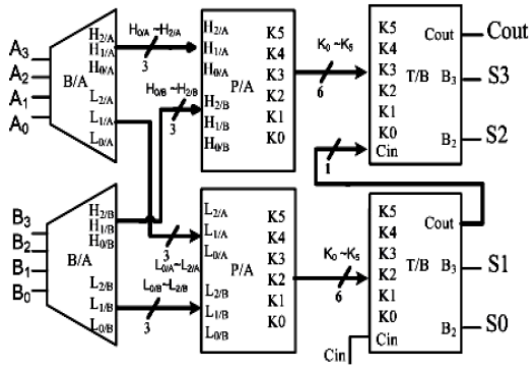


Figure 5. Block diagram of the 4-bit abacus adder[10].

Block diagram of radix-4 abacus adder is depicted in figure 5. As shown in figure 5, the function of radix-4 abacus adder is completed in four phases. In first phase binary input is converted into abacus output known as B/A(binary to abacus) module. In second phase, parallel additon takes place known as P/A(parallel addition module). In the last (third) phase thermometric input converted into binary output known as T/B(thermometric to binary) transformation module.

V. IMPLEMENTATION OF PROPOSED DESIGN AND SIMULATION RESULTS

The 4x4 Vedic multiplier is design using the Chinese abacus methodology. It may simply use three 4-bit abacus adders, shown in figure 3. We design all component of multiplier using reversible logic gates.

The digital logic implementation of the 2X2 Urdhva Tiryakbhayam multiplier using the conventional logic gates [11] is as shown in figure 6. The expressions for the four output bits are given below figure 6. The block diagram of reversible logic gate implementation of 2x2 UT multiplier is as shown in figure 7. This design does not consider the fanouts.

The circuit requires a total of six reversible logic gates out of which five are Peres gates and remaining one is the Feynman Gate. The quantum cost of the 2X2 Urdhva Tiryakbhayam Multiplier is enumerated to be 21. The number of garbage outputs is 9 and number of constant inputs is 4.

The Reversible 4X4 Urdhva Tiryakbhayam Multiplier design emanates from the 2X2 multiplier. The block diagram of the 4X4 Vedic Multiplier is presented in the figure 3. It consists of four 2X2 multipliers, each of which processes four bits as inputs; two bits from the multiplicand and two bits from the multiplier. The lower two bits of the output of the first 2X2 multiplier are entrapped as the lowest two bits of the final result of multiplication. Two zeros are concatenated with the upper two bits and given as input to the four bit abacus adder.

The other four input bits for the abacus adder are obtained from the second 2X2 multiplier. Likewise, the outputs of the third and the terminal 2X2 multipliers are given as inputs to the second four bit abacus adder. The outputs of these four bit adders are in turn 5 bits each, which need to be summed up. This is done by a five bit abacus adder which generates a six bit output. These six bits form the upper bits of the final result.

The Chinese abacus adder is consummated (realized) using the NFT and F2G Gate. Single NFT Full Adder (SNFA**)** is a Fault Tolerant full adder circuit which consists of one New Fault Tolerant (NFT) gate and three Feynman Double (F2G) gates where the quantum cost is 11 and the total number of garbage outputs is 3.

This design also does not take into consideration the fanouts of the gates. For this design the quantum cost is computed to be 162, the total number of gates used will be 36, the number of garbage outputs will be 62 and the number of constant inputs will be 29.

Copyright to IJARCCE
DOI 10.17148/IJARCCE.2015.4418

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

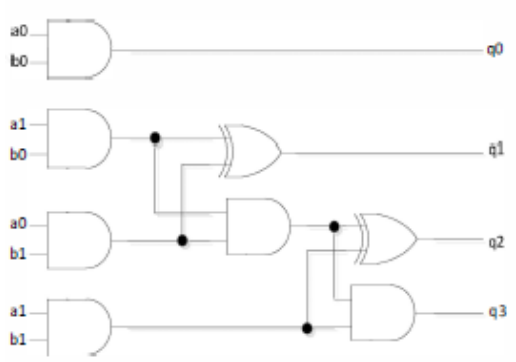


Figure 6: Conventinal Logic Implementation of 2x2 UT multiplier

q0 = a0.b0
ql= (a1.b0) xor (a0.bl)
q2= (a0.al.b0.bl) xor (al.bl)
q3= a0.al.b0.bl

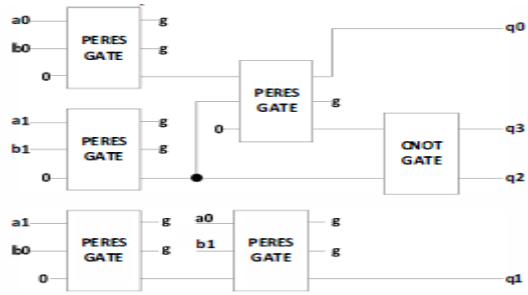


Figure 7: Reversible logic gate implementation of 2x2 UT multiplier.

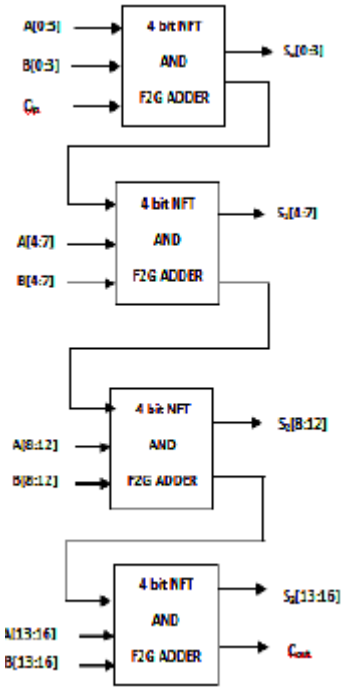


Figure 8: Reversible logic gate implementation of Chinese Abacus Adder.

Simulation of the proposed work is done using Modelsim SE 6.5. Synthesis and Implementation is done using Xilinx Spartan-3E FPGA Board of device family xc3s50-4-tq144. The subsequent figures show the simulation, RTL and Technical Schematic of the multiplier. Results show that 4x4 UT vedic multiplier designed in this work using reversible logic gates, consumes less area and has high speed as compared to 4x4 UT vedic multiplier designed using conventional logic gates. Results also show that 4x4 UT vedic multiplier designed using reversible logic gates is 23% faster and takes less area than 4x4 UT vedic multiplier design without using reversible logic gates

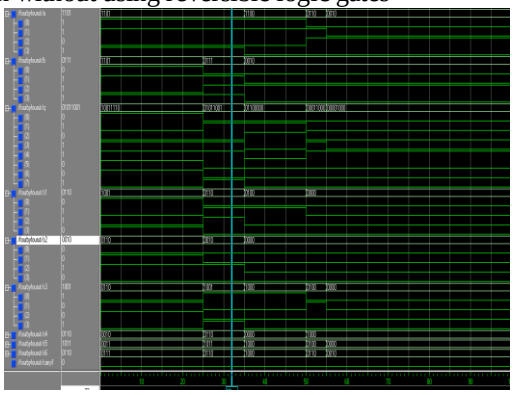


Figure 9: Simulation result for 4x4 bit Vedic multiplier designed using reversible Logic gates

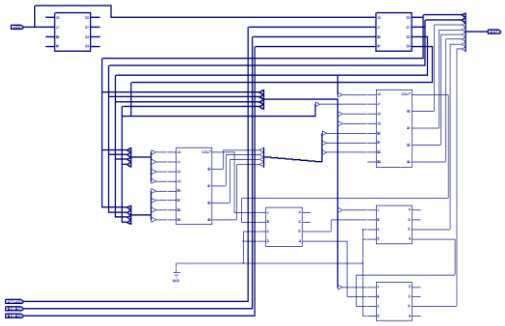


Figure 10: RTL view of 4x4 bit UT Vedic multiplier designed using reversible logic gates.

TABLE I: SIMULTAION RESULTS OF DEVICE UTILIZATION

Methods	**UT Vedic multiplier without reversible logic gates**	**UT Vedic multiplier with reversible logic gates**
Number of slices	36 out of 768 (4%)	28 out of 768 (3%)
Number of 4 input LUTs	65 out of 1536 (4%)	49 out 1536 (3%)
Number of bonded IOBs	17 out of 97 (17%)	16 out of 97 (17%)
Net propagation delay	20.705 ns	15.830 ns
Frequency	48.297MHz	63.17MHz

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 4, April 2015

VI. CONCLUSION

The use of Vedic multiplier approach results in a competitive technique as compared respect to conventional fast multiplier. The simulation results show that this approach of Vedic multiplier using Chinese Abacus adder is very efficient for low-power, high-speed applications. This architecture is also easy for pipeline implementation. Another advantage is that the methodology may reduce the number of ripple carries and partial product generation in many steps. It can be inferred that Vedic multiplier using Chinese Abacus adder is quite efficient as compared to the conventional multiplier, for multiplication operation.

REFERENCES

[1]. R. Feynman, "Quantum Mechanical Computers," Optics News, Vol.1l, pp. 11-20, 1985.
[2]. Peres, "Reversible logic and quantum computers", Phys. Rev. A 32 (1985) 3266-3276.
[3]. E. Fredkin and T. Toffoli,"Conservative Logic", Int'l 1 Theoretical Physics Vo121, pp.219-253, 1982.
[4]. Shams, M., M. Haghparast and K. Navi, Novel reversible multiplier circuits in nanotechnology. World Appl. Sci. J., 3(5): 806-810.
[5]. Shailja Shukla, Tarun Verma and Rita Jain, "Design of 16 Bit Carry Look Ahead Adder Using Reversible Logic", International Journal of Electrical, Electronics and Computer Engineering 3(1): 83-89(2014).
[6]. Franco Maloberti and Chen Gang, "The Chinese Abacus method: can we use it for digital arithmetic", Proceedings of the 8th Great Lakes Symposium on VLSI, pp. 192 - 195 19-21 Feb. 1998.
[7]. Franco Maloberti and Chen Gang, "Use of the Chinese Abacus method for digital arithmetic functions", Proceedings of the 1998 IEEE International Symposium on Circuits and Systems, Vol. 5, pp. 213 - 216, vol.5, 31 May - 3 June 1998.
[8]. Franco Maloberti and Chen Gang, "Performing Arithmetic Functions with the Chinese Abacus Approach," IEEE Transactions on circuits and systems-II: Analog and digital signal processing, vol. 46, no. 12, pp. 1512- 1515, Dec. 1999.
[9]. Andreev, B.D., Titlebaum, E., Friedman, E.G., "Tapered Transmission Gate Chains for Improved Carry Propagation,"MWSCAS-2002. The 45th Midwest Symposium on Circuits and Systems, Volume 3, pp. 449 - 452, Aug. 4-7, 2002.
[10]. Behrooz Parhami "Computer Arithmetic-Algorithms and Hardware Designs", Oxford University Press, Inc (2000).
[11]. G Ganesh Kumar and V Charishma, "Design of high speed vedic multiplier using vedic mathematics techniques", ltn'l J. of Scientific and Research Publications, Vol. 2 Issue 3 March 2012.
[12]. Maloberti, F., Chen Gang, "Functions with the Chinese Abacus Approach", IEEE Transaction on circuits and systems-II: Analog and digital signal processing, vol. 46, no. 12, pp. 1512 – 1515, Dec.1999.
[13]. Shu-Chung Yi, Kun-Tse Lee, Jin-Jia Chen, Chien-Hung Lin, Chuen-Ching Wang, "The new architecture of radix-4 Chinese abacus adder", Proceedings of the 36th International Symposium on Multiple-Valued Logic (ISMVL'06) 2006 IEEE.
[14]. Shu-Chung Yi, et al.,"Comments on Performing Arithmetic Functions with the Chinese Abacus Approach", IEEE Transaction on circuits and systems-II.
[15]. Rakshith Saligram and Rakshith T.R. "Design of Reversible Multipliers for linear filtering Applications in DSP", International Journal of VLSI Design and Communication systems, Dec-2012.
[16]. .Harpreet Singh Dhillon Abhijit Mitra, "A Digital Multiplier Architecture using Urdhava Tiryabhyam Sutra of Vedic Mathematics", International Conference on Design & Technology of Integrated Systems in Nanoscale Era, 2008.
[17]. Dr. Malti Bansal, Diksha Ruhela, "High Speed & Area Efficient Vedic Multiplier using Adiabatic Logic", Journal of Basic and Applied Engineering Research Volume 1, Number 11; October-December 2014 pp. 14-17.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Credit Card Fraud Detection Using Advanced Combination Heuristic and Bayes' Theorem

Sahil Hak[1], Suraj Singh[2], Varun Purohit[3]

Software Engineering Undergraduate, Department of Computer Engineering, Delhi Technological University, Delhi, India[1,2,3]

ABSTRACT: In today's increasingly electronic society and with the rapid advances of electronic commerce on the Internet, the use of credit card transactions have become the standard for Internet and Web-based e-commerce. In this paper, we present a credit card fraud detection model which takes into account present as well as the past behaviour. The detection model comprises of card validation via Luhn's algorithm, two initial probability assignments based on address mismatch and spending pattern, an advanced combination heuristic, spending-pattern database and Bayes' theorem. Advanced combination heuristic is an improvement that eliminates the conflict of existing Dempster-Shafer theory.

KEYWORDS: CCFDS; credit-card; Luhn's algorithm; Longest Common Subsequence; DBSCAN; advanced combination heuristic; Bayes' theorem

I. INTRODUCTION

With rapid advancement of e-commerce, use of credit cards for purchases has exponentially increased. Unfortunately, fraudulent use of credit cards has also become a source of crime. Credit card fraud is a wide-ranging term for theft and fraud committed using or involving a payment card, such as a credit card or debit card, as a fraudulent source of funds in a transaction. Credit card fraud is also an adjunct to identity theft. According to the United States Federal Trade Commission, while identity theft had been holding steady for the last few years, it saw a 21 percent increase in 2008. However, credit card fraud, that crime which most people associate with ID theft, decreased as a percentage of all ID theft complaints for the sixth year in a row.[1]

According to a study conducted by ACI Worldwide and Aite Group released in June 2014, about 41 per cent of card users in India have experienced card fraud during 2012-2014. Study also highlighted that the highest rate of fraud on prepaid cards is experienced by consumers in India at 18 per cent.[2]

Financial institutions employ various fraud prevention models for tackling this problem. But fraudsters are adaptive, and given time, they devise several ways to intrude such protective models. Despite the best efforts of the financial institutions, law enforcement agencies and the government, credit card fraud continues to rise. Fraudsters nowadays may constitute of a very inventive, intellect and fast moving fraternity. Several techniques for the detection of credit card fraud have been proposed in the last few years. We briefly review some of them in Section II and propose our model in further sections.

II. RELATED WORK

Various existing fraud techniques majorly explore decision trees, genetic algorithms, clustering techniques and neural networks.

Recently, Syeda et al.[3] have suggested the use of parallel granular neural networks for speeding up the data mining and knowledge discovery process. Maes et al.[4] have outlined an automated credit card fraud detection system by Artificial Neural Network - ANN as well as Bayesian belief networks - BBN. They show that BBN gives better results related to fraud detection and the training period is faster whereas the actual detection process is substantially faster

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

with ANN. The neural network based methods are, in general, fast but not so accurate an retraining the neural networks is quite taxing.

Chen et al.[5] propose a method in which an online questionnaire is used to collect questionnaire-responded transaction (QRT) data of users. Further it uses a support vector machine (SVM) trained with this data and the QRT models are used to predict new transactions. Chen et al.[6] have recently presented a personalized approach for credit card fraud detection that employs both SVM and ANN. It tries to prevent fraud for users even without any transaction data. However, these systems are not fully automated and depend on the user's expertise level.

Researchers have applied data mining for credit card fraud detection. Chan et al.[7] divide a large set of transactions into smaller subsets and then apply distributed data mining for building models of user behavior. The resultant base models are then combined to generate a meta-classifier for improving detection accuracy. Brause et al.[8] have explored the possibility of combining advanced data mining techniques and neural networks to obtain high fraud coverage along with a low false alarm rate. Use of data mining is also elaborated in the work by Chiu and Tsai[9]. They consider web services for data exchange among banks. A fraud pattern mining (FPM) algorithm has been developed for mining fraud association rules which give information regarding the features that exist in fraud transactions. Banks enhance their original fraud detection systems by using the new fraud patterns to prevent attacks. While data mining techniques are relatively accurate, they are inherently slow.

Aleskerov et al.[10] present CARDWATCH, a database mining system used for credit card fraud detection. The system, based on a neural learning module, provides an interface to a variety of commercial databases. Kim and Kim have identified skewed distribution of data and mix of legitimate and fraudulent transactions as the two main reasons for the complexity of credit card fraud detection[11]. Based on this observation, they use fraud density of real transaction data as a confidence value and generate the weighted fraud score to reduce the number of wrong detections.

Credit card fraud detection by S Panigrahi et al.[12] suggests a fusion approach using Dempster-Shafer theory and Bayesian learning. Results from rule based filters are used and combined them using Dempster - Shafer adder. Later, these beliefs are strengthened by transaction history database and Bayesian learning.

Dempster has given a rule to combine evidences coming from different independent sources. However, Dempster's rule of combination has been criticized as sometimes it gives some illogical results. Therefore the above proposed model is susceptible to failures.

We use a better and more elastic combination rule by Tazid Ali et al.[13] discussed in further sections that eliminates the conflict problem of Dempster-Shafer theory.

III. PROPOSED ALGORITHM

The proposed model comprised of six steps. Firstly, Luhn's Test is used to validate card numbers. Then, two rules ie. Address Mismatch and Degree of Outlierness are used to analyze the deviation of each incoming transaction from the normal profile of cardholder. These two steps compute initial beliefs. The initial belief values are combined to obtain an overall belief by applying Advanced Combination Heuristic in step four. Step five looks into the spending history to extract characteristic information about genuine and fraud transactions. The overall belief is further strengthened or weakened in the final step using Bayes' Theorem, followed by recombination of the calculated probability with initial belief of fraud using advanced combination heuristic.

Step 1 : Luhn's Algorithm

Following standard algorithm is used to validate credit card numbers,[14]

1. Reverse the order of the digits in the number.
2. Take the first, third, ... and every other odd digit in the reversed digits and sum them to form the partial sum S1.
3. Taking the second, fourth and every other even digit in the reversed digits. Multiply each digit by two and sum the digits if the answer is greater than nine to form partial sums for the even digits.
4. Sum the partial sums of the even digits to form S2.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

5. If S1+S2 ends in zero,then the original number is in the form of a valid credit card number as verified by the Luhn test.

For example, if the trial number is 49927398716,

1. Reverse the digits:
   61789372994
2. Sum the odd digits:
   6 + 7 + 9 + 7 + 9 + 4 = 42 = s1
3. The even digits:
   1, 8, 3, 2, 9
   Two times each even digit:
   2, 16, 6, 4, 18
   Sum the digits of each multiplication:
   2, 7, 6, 4, 9
4. Sum the last:
   2 + 7 + 6 + 4 + 9 = 28 = s2
5. S1+S2 = 70 which ends in zero which means that 49927398716 passes the Luhn's test.

Step 2 : Address Matching

The rule checks whether the Billing Address and Shipping Address match or not. This check does not guarantee whether a transaction is fraud or genuine. But if the two addresses match,the transaction can be classified as genuine with a high probability. Else,the transaction is labelled as suspect.

Suppose the address has the format: Road Number, House Number, Street Address, City Code, Pin Code. Let the billing address be: Road Number 76, House Number 24, Wall Street, New Jersey, 123214. If the billing house number is 76, and the user enters house number 73 as the shipping house number, the 2 numbers don't match. This check can be directly performed by comparing the equality of the 2 numbers.

But if the billing street address is 'Wall Street' and the shipping street address is 'Wll Strt',there is scope for matching, since the user might have entered a shorter form of the address. For this check to be successful, we have to initially store the billing address in it's longest form. For example, Billing Address should be stored as 'Wall Street' and not 'Wll Strt'.

1. Count the number of characters in Billing street Address and shipping street address. Label them as C1 and C2. If length of shipping street address is greater than 60% of the length of billing street address, we can proceed to step 2. Otherwise, we can assume that the addresses do not match.
2. Find whether the number of words of billing street address is equal to number of words in shipping street address. If equal, we can proceed to step 3. Else, the addresses do not match.
3. Find whether the first letter of every word in shipping street address is same as the first letter of corresponding word in billing street address. If yes, we can proceed to step 4. Else, addresses do not match.
4. Find the longest common subsequence of the two street address. If the longest common subsequence is same as shipping street address,addresses match. Else, they don't.[15]

Step 3 : Outlier Detection

A customer usually carries out similar types of transactions in terms of amount, which can be visualized as part of a cluster. Since a fraudster is likely to deviate from the customer's profile, his transactions can be detected as exceptions to the cluster – a process known as outlier detection.

We have used DBSCAN[16] (Density Based Spatial Clustering of Application with Noise) to generate clusters, using transaction amount as attribute. Any incoming transaction amount,that does not belong to any cluster is detected as outlier. Such an observation gives evidence that the transaction could be fraudulent. We measure the extent of deviation of an incoming transaction by its degree of outlierness. If the average distance of the amount p of an outlier transaction T from the set of existing clusters is x, then its degree of outlierness d_outlier is given by,

d_outlier,

$1 - \frac{E}{x}$ If Neighbouring points of P < MinPts

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

0 Otherwise

MinPts: Minimum number of points required in the epsilon-neighborhood of each point to form a cluster.
E(epsilon): Maximum radius of the neighborhood.

Step 4 : Advanced Combination Heuristic

More often, it is seen that available information is interpreted in probabilistic sense because probability theory is a very strong and well established mathematical tool to deal with objective uncertainty (i.e., uncertainty arises from heterogeneity or the random character of natural processes). However, it is clear that not all available information, data or model parameters are affected by objective uncertainty (i.e., nature of the data, information or parameters are random) and can be handled by traditional probability theory. Imprecision may occur due to scarce or incomplete information or data, measurement error or data obtain from expert judgment or subjective interpretation of available data or information. Thus, model parameters, data may be affected by subjective uncertainty. Traditional probability theory is inappropriate to represent subjective uncertainty (i.e., uncertainty arises from the partial character of our knowledge of the natural world). To overcome the limitation of probabilistic method, Dempster put forward a theory and now it is known as evidence theory or Dempster-Shefer theory (1976).

For the credit card fraud detection problem, Demspter-Shafer Theory is more relevant as compared to other fusion methods since it introduces a third alternative: ‘‘unknown”, along with the measure of confidence in each of the alternatives. It provides a rule for computing the confidence measures of three states of knowledge: fraud, fraud and suspicious (unknown) based on data from new as well as old evidence. Furthermore, in DST, evidence can be associated with multiple possible events unlike traditional probability theory where evidence is associated with only one event. As a result, evidence can be more meaningful at a higher level of abstraction. Hence, we use Dempster–Shafer theory for combining evidences for this problem. However, one of the shortcomings of DST is that, for evidences with a high degree of conflict, the modeling may not be accurate.

Given two mass functions m_1 and m_2, Dempster-Shafer theory also provides a combination rule for combining them, which is defined as follows,

$$m(h) = m_1(h) \oplus m_2(h) = \frac{\sum_{x \cap y = h} m1(x) * m2(y)}{1 - \sum_{x \cap y = \phi} m1(x) * m2(y)}$$

The above rule of combination produces unreasonable results as demonstrated in the following example, Deqiang et al. 2008 [17]. Zadeh had given a compelling example of erroneous result found by using D-S theory. Suppose two doctors examine a patient and agree that it suffers from either meningitis (M), contusion (C) or brain tumor (T). Thus frame of discernment = {M, C, T}. Assume that the doctors agree in their low expectation of a tumor, but disagree in likely cause and provide following diagnosis:

$$m_1(M) = 0.99,\ m_1(T) = 0.01$$
$$m_2(C) = 0.99,\ m_2(T) = 0.01$$

Based on Dempster rule of combination, we get the unexpected final conclusion of m(T) = 1.
It means that the patient suffers with certainty from brain tumor. This unexpected result arises from the fact that the two bodies of evidence (doctors) agree that the patient most likely does not suffer from tumor but are in almost full contradiction for the other causes of the disease.

Using the advanced combination heuristic[13],we can avoid the conflict of Dempster of Dempster Shafer theory.
Let (F, m_1) and (F, m_2) be two body of evidences given by two independent experts. Suppose A_1, A_2, ….., A_n be the focal elements. We will construct a combine body of evidence by assigning new bpa to each of the focal elements A_1, A_2, ….., A_n. For each A_i we associate a value $m'(A_i)$ as,

$$m'(A_i) = m_1(A_i) + m_2(A_i) - m_1(A_i) * m_1(A_i)$$

This value can be interpreted as the combined evidence in support of the focal elements A_i . The motivation behind this formulation is the probability rule for union of two events. It also resembles that algebraic sum of fuzzy sets. Then the quantity $m'(A_i) / \{1 - m'(A_i)\}$ can be interpreted as the odds ratio in favour of A_i, which can also be considered as evidence in support of A_i. Since we will deal with conflict situation in evidence theory the denominator of the above expression may be 0. So, we replace the denominator by $1 + \{1 - m'(A_i)\}$, which gives,

$$m'(A_i) = \frac{1 - (1 - m1(Ai)) * (1 - m2(Ai))}{1 + (1 - m1(Ai)) * (1 - m2(Ai))}$$

Finally we assign bpa of A_1 as,

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

$$m(A_i) = \frac{m'(A_i)}{\sum_n m'(A_i)}$$

Using our proposed combined rule we solved the compelling example given by Zadeh in which the frame of discernment is Θ = {M,C, T} and corresponding bpas are $m_1(M) = 0.99$, $m_1(T) = 0.01$ and $m_2(C) = 0.99$, $m_2(T) = 0.01$. Our method gives the bpa as m(M) = 0.494745, m(C) = 0.49745 and m(T) = 0.0051 which is more reliable than Dempster's rule of combination.

Step 5 : Spending Pattern Database

It comprises of genuine Transaction Record(for individual customers from their past behaviour) and Fraud Transaction Record(from different types of past fraud data).We represent each history transaction by set of attributes containing information like card number, transaction amount and time since last purchase.

To capture the frequency of card use, we consider the time gap between successive transactions on the same card. The transaction gap is divided into Ten mutually exclusive and exhaustive events – D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}. Occurrence of each event depends on the time since last purchase (transaction gap) on any particular card. Let Y represents the time gap. The event D_1 is defined as the occurrence of a transaction on the same card within 15 hours of the last transaction which can be represented as,

D_1 = True | 0 < Y <=15
D_2 = True | 15 < Y <=30
D_3 = True | 30 < Y <=45
D_4 = True | 45 < Y <=60
D_5 = True | 60 < Y <=75
D_6 = True | 75 < Y <=90
D_7 = True | 90 < Y <=105
D_8 = True | 105< Y <=120
D_9 = True | 120< Y <=135
D_{10} = True | Y>135

We next compute $P(D_i|h)$ and $P(D_i|h')$ from the Fraud Transaction Record and the Good Transaction Record, respectively. $P(D_i|h)$ measures the probability of occurrence of D_i given that a transaction is originating from a fraudster and $P(D_i|h')$ measures the probability of occurrence of D_i given that it is genuine.

$P(D_i|h)$ = (Number of Occurrences of D_i in FTR) / (Number of transactions in FTR)
$P(D_i|h')$ = (Number of Occurrences of D_i on card k in GTR) / (Number of transactions on card k in GTR)
FTR : Fraud Transaction Record
GTR : Genuine Transaction Record
$P(D_i) = P(D_i|h) * P(h) + P(D_i|h') * P(h')$

Step 6 : Bayes' Theorem

The idea of belief revision is that, whenever new information becomes available, it may require updating of prior beliefs. Bayes' rule gives the mathematical formula for belief revision, which can be expressed as,

$$P(D_i|h) = \frac{P(D_i|h) * P(h)}{P(D_i)}$$

From Step 5,

$$P(D_i|h) = \frac{P(D_i|h) * P(h)}{P(D_i|h) * P(h) + P(D_i|h') * P(h')}$$

Now, P(h) and $P(D_i|h)$ are combined using advanced combination heuristic discussed in step 4, to get the final belief of fraud. This final belief is now compared against the threshold values to determine whether the transaction is genuine, fraud or suspicious.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Methodology

The card is first validated by Luhn's algorithm is step 1. Incoming transaction is first handled by two rules, address matching and outlier detection. Probability values from these two rules are combined using advanced combination heuristics to get the initial belief of fraud P(h) for the transaction. Ω_L and Ω_U are the upper and lower threshold values for transaction classification into suspicious, genuine and fraud. These values can be learned over time or can be manually fed if sufficient information is available with the bank. If P(h) is less than Ω_L, the transaction is considered to be genuine and is approved. If P(h) is greater than Ω_U, the transaction is considered to be fraudulent and overhead safety precautions are taken. In case P(h) is in between Ω_L and Ω_U, card is labeled as suspicious.

In case the transaction is found to be suspicious, it is inserted in the suspect table. Credit Card Fraud Detection System determines which event D has occurred out of the ten Di's explained in the step 6 above, and retrieves the corresponding P(E|h) and P(E|h') values from the tables FTR (Fraud Transaction Record) and GTR (Good Transaction Record), respectively. The beliefs P(h|E) and P(h'|E) are next computed using Bayes' theorem. The initial belief of fraud is further strengthened by combining these probabilities with initial belief of fraud using advanced combination heuristic. Now accordingly, the transaction is treated upon using threshold values Ω_L, the lower threshold and Ω_U, the upper threshold. Fraudulent transactions are updated in Fraud Transaction Record and genuine transactions are updated in Good Transaction Record.

IV. SIMULATION AND RESULTS

Basic Probability Assignments,

1. Address Matching

If address mismatch occurs then there is a high probability that it is a fraud transaction and low probability that it is a genuine transaction. Alternatively, if the addresses match, then there is a high probability that it is a genuine transaction and low probability that it is a fraud transaction.

If address mismatch occurs,we assume the following basic probability assignments,

$m_1(h) = 0.6$
$m_1(h') = 0$
$m_1(U) = 0.4$

If $m_1(h)$ is set too high, the probability that a transaction is detected as fraudulent will go up. This improves the detection rate but also raises the number of false alarms. Similarly, if $m_1(U)$ is set high, the number of suspicious transactions goes up, increasing the number of misses. Thus, the values 0.6 and 0.4 have been chosen to maintain a balance.

If addresses match,we assume the following probability assignments,

$m_1(h) = 0$
$m_1(h') = 0.6$
$m_1(U) = 0.4$

2. Outlier Detection

For a transaction detected as an outlier, we make the following basic probability assignments using the degree of outlierness,

$m_2(h) = 1 - (E/x)$ (E , x hold meanings as discussed in step 3)
$m_2(h') = 0$
$m_2(U) = 1 - (1 - (E/x))$

The test run was conducted with the tailor made data which covered all the varied possibilities. Wide ranges of both the band pass filters of address mismatch and outlier detection were taken into consideration. Seven of the many sample runs of different categories of transactions were selected and are shown below[Table I]

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

S.No	Address Mismatch	Degree of Outlierness	Initial belief of fraud	Category of Transaction
1	No	0.86	0.736844	Fraud
2	Yes	0.77	0.440058	Suspicious
3	No	0.46	0.531902	Suspicious
4	No	0.12	0.356390	Suspicious
5	Yes	0.15	0.060305	Genuine
6	Yes	0.26	0.114321	Genuine
7	Yes	0.46	0.230716	Genuine

Table I : Test Run for various category of transactions. Address matching from Step 2. Degree of Outlierness from Step 3. Initial belief of fraud from advanced combination heuristics from Step 4. Category of transactions are fraud for score greater than 0.7, genuine for score less than 0.3 and suspicious if the score is in between.

As mentioned in part III of this paper, the algorithm needs to learn threshold values of fraud, suspicious and genuine transactions or these values can be fed directly, if are available with the bank. We assumed initial belief of fraud less than 0.3 to be genuine,and initial belief of fraud greater than 0.7 to be fraud.

In Table I, various cases are tested, both for address match and address mismatch (Step 2). Degree of outlierness is computed next using DBSCAN algorithm (Step 3). The probabilities from these two rules are combined using advanced combination heuristic (Step 4) to get the initial belief of fraud. In Case 1, initial belief of fraud is 0.736844,which is greater than 0.7, so the transaction is marked as fraudulent. In Case 5, 6 and 7, initial belief of fraud is 0.060305, 0.114321 and 0.230716 respectively. They all are less than 0.3, so the transactions are marked as genuine.

Scores of transaction 2,3 and 4 are 0.440058, 0.531903 and 0.356390 respectively. They all lie in the suspicious zone (between 0.3 and 0.7). Therefore, they undergo one more round where-in we find P(E/h) and P(E/h') using the spending pattern database and then compute the bayesian probabilities P(h/E) and P(h'/E). The bayesian probabilities are then combined with initial belief of fraud using advanced combination heuristic to get the final belief of fraud. This gives us the reevaluated category of transaction. We apply the same concept of thresholds here as well. This is shown in the table[Table II] below.

S.No	P(E/h)	P(E/h')	P(h/E)	P(h'/E)	Final belief of fraud	Reevaluated Category of Transaction
2	0.54	0.58	0.422	0.577	0.42530	Suspicious
3	0.28	0.61	0.341	0.658	0.43252	Suspicious
4	0.76	0.47	0.471	0.528	0.40925	Suspicious

Table II : Further round for suspicious transactions. P(E/h), P(E/h') from Step 5. P(h/E), P(h'/E) from Step 6. Final belief of fraud computed using advanced combination heuristic (Step 4).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

REFERENCES

1. Federal Trade Commission, "Consumer Sentinel Network Data Book: January - December 2008", 26 February 2009.
2. ACI Worldwide, "Global Consumers : Losing Confidence in the Battle Against Fraud", June 2014.
3. M. Syeda, Y.Q. Zhang, Y. Pan, "Parallel granular neural networks for fast credit card fraud detection", Proceedings of the IEEE International Conference on Fuzzy Systems, 2002, pp. 572–577.
4. S. Maes, K. Tuyls, B. Vanschoenwinkel, B. Manderick, "Credit card fraud detection using Bayesian and neural networks", Proceedings of the First International NAISO Congress on Neuro Fuzzy Technologies, 2002.
5. R.C. Chen, M.L. Chiu, Y.L. Huang, L.T. Chen, "Detecting credit card fraud by using questionnaire-responded transaction model based on support vector machines", Proceedings of the Fifth International Conference on Intelligent Data Engineering and Automated Learning, vol. 3177, October 2004, pp. 800–806.
6. R.C. Chen, S.T. Luo, X. Liang, V.C.S. Lee, "Personalized approach based on SVM and ANN for detecting credit card fraud", Proceedings of the IEEE International Conference on Neural Networks and Brain, October 2005, pp.810–815.
7. P.K. Chan, W. Fan, A.L. Prodromidis, S.J. Stolfo, "Distributed data mining in credit card fraud detection", Proceedings of the IEEE Intelligent Systems, 1999, pp.67–74.
8. R. Brause, T. Langsdorf, M. Hepp, "Neural data mining for credit card fraud detection", Proceedings of the International Conference on Tools with Artificial Intelligence, 1999, pp. 103–106.
9. C. Chiu, C. Tsai, "A web services-based collaborative scheme for credit card fraud detection", Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service, 2004, pp. 177–181.
10. E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection", Computational Intelligence for Financial Eng., pp. 220-226, 1997.
11. M.J. Kim and T.S. Kim, "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection," Proc. Int'l Conf. Intelligent Data Eng. and Automated Learning, pp. 378-383, 2002.
12. Suvasini Panigrahi, Amlan Kundu, Shamik Sural, A.K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning" , Information Fusion 10 (2009) 354–363.
13. Tazid Ali, Palash Dutta, Hrishikesh Boruah, "A New Combination Rule for Conflict Problem of Dempster-Shafer Evidence Theory", International Journal of Energy, Information and Communications, Vol. 3, Issue 1, February, 2012.
14. Hans P Luhn, "Computer for Verifying Numbers", US Patent 2,950,048, August 23, 1960 (Public Domain)
15. Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest and Clifford Stein (2001). "15.4 - *Introduction to Algorithms* (2nd ed.)", MIT Press and McGraw-Hil,. pp. 350–355, ISBN 0-262-53196-8.
16. Ester, Martin; Kriegel, Hans-Peter; Sander, Jörg; Xu, Xiaowei (1996). Simoudis, Evangelos; Han, Jiawei; Fayyad, Usama M., eds.,"A density-based algorithm for discovering clusters in large spatial databases with noise", Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press. pp. 226–231. ISBN 1-57735-004-9.CiteSeerX: 10.1.1.71.1980.
17. H. Deqiang, H. Chongzhao, and Y. Yi, "A modified evidence combination approach based on ambiguity measure", Proceedings of the 11th International Conference on Information fusion, (2008), pp.1-6.

Volume 3, Issue 1 (2015) 189-196 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Effect of Single and Multiple-Pass Friction Stir Processing on Microstructure, Hardness and Tensile Properties of a 99.99% Cu with Carbon Nano Tubes

V. Jeganathan Arulmoni, Ranganath M. S., R. S. Mishra
Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

Article Info
Article history:
Received 20 February 2015
Received in revised form
25 February 2015
Accepted 28 February 2015
Available online 15 March 2015

Keywords
Microstructure,
Hardnes,
Ttensile Properties,
Friction Stir Processing,
Copper,
Carbon Nano Tubes

Abstract
Friction stir processing (FSP), a relatively new solid-state microstructural modification technique can also be viewed as a severe plastic deformation technique because of the high processing strain involved. It is also unique from the viewpoint of its applicability to a localized region. Here the states of development of FSP for processing of Copper with carbon nanotubes are addressed. This paper investigates the parameters affecting the friction stir processed copper with carbon nano tubes and enhancement of the microstructure, hardness and tensile properties of the composite material. The behaviour of Copper with carbon nano tubes has been studied with single pass, double passes and triple passes. The SEM results are also presented for each case.

1. Introduction

Friction stir processing (FSP), a relatively new solid-state microstructural modification technique can also viewed as a severe plastic deformation technique because of the high processing strain involved. It is also unique from the viewpoint of its applicability to a localized region. FSP was developed for microstructural modification of metallic materials. In FSP, a rotating tool is plunged into a material and high plastic deformation is produced. FSP is used to enhance ductility, induces super plasticity and improve corrosion resistance properties. Dynamic recrystallization of the deformed zone forms an ultrafine-grained structure. FSP has been successfully applied to various cast aluminium and magnesium and copper alloys to eliminate casting defects and there by improve their mechanical properties [24, 25]. Two modes of metal transfer during friction stir processing have been discussed [15]. The first mode of metal transfer is generated between the tool shoulder and the plate and takes place as layer-by-layer deposition of metal one over the other. The second mode of metal transfer is generated by the extrusion of metal around the tool pin, when it reaches a state of sufficient plasticity. Metal transfer, generated between the tool shoulder and the plate, plays an important role in influencing the mechanical properties during friction stir process [29]. Modes of metal transfer are clearly visible in the microstructure characteristics, but they are not too distinct in macrostructure of most processed samples. Friction stir processing can be applied as a single-pass for processing a small area. For large engineering components in which the contact areas are relatively large, single pass FSP may not be adequate. Multi-pass FSP with a certain level of overlap between the successive passes is required for large contact areas. For both single and multi-pass processes, it is important to assess the microstructural evolution and its influence on the mechanical properties [2, 7, 17, and 23]

Copper is a mostly used industrial and functional metal for various thermal, electrical and electronic applications, i.e. electronic packaging, electrical contacts and resistance welding electrodes. This is because of good thermal and electrical conductivity, high plasticity and excellent resistance to corrosion and oxidation. Nevertheless, low mechanical strength and poor wear resistance limit its applications [3, 10, 16, 26, and 27]. Carbon nano tubes are unique tubular structures of nanometer diameter and large length/diameter ratio. The nanotubes may consist of one up to tens and hundreds of concentric shells of carbons with adjacent shells separation of _0.34 nm. The carbon network of the shells is closely related to the honeycomb arrangement of the carbon atoms in the graphite sheets. The amazing mechanical and electronic properties of the nanotubes stem in their quasi-one dimensional (1D) structure and the graphite-like arrangement of the carbon atoms in the shells. Thus, the nanotubes have high Young's modulus and tensile strength, which makes them preferable for composite materials with improved mechanical properties. The nanotubes can be metallic or semiconducting depending on their structural parameters. This opens the ways for application of the nanotubes as central elements in electronic devices including field-effect transistors (FET), single-electron transistors and rectifying diodes. Possibilities for using of the nanotubes as high-capacity hydrogen storage media were also considered for experiment purpose [26].

Here the states of development of FSP for processing of Copper with carbon nano tubes are addressed. This paper investigates the parameters affecting the friction stir processed copper with carbon nanotubes and enhancement of the microstructure, hardness and tensile properties of the composite material. The behaviour of Copper with carbon nano tubes has been studied with single pass, double passes and triple passes The SEM results are also presented for each case.

Corresponding Author,
E-mail address: v.jeganathan.dce@gmail.com
All rights reserved: http://www.ijari.org

2. Literature Review

Bahram A. Khiyavi, et al [1] produced copper reinforced metal matrix composite (MMC) using micron sized chromium particles via friction stir processing (FSP) in order to study effects of adding Cr particles to copper based matrix by FSP. Microstructures, micro hardness and wear properties were studied in order to evaluate the microstructures and mechanical properties of fabricated composites. H.R. Akramifard et al [4] in their investigation, pure Cu sheets were reinforced with 25μm SiC particles to fabricate a composite surface layer by friction stir processing (FSP). In order to improve distribution of reinforcing SiC particles, a net of holes were designed by drill on the surface of pure Cu sheet. H. Sarmadi, et al [5] focused on friction stir processing (FSP) used to produce copper–graphite surface composites. Five tools with different pin profile were employed in order to achieve a comprehensive dispersion. Results showed that the tool with rectangular pin give rise to a better dispersion of graphite particles. Furthermore, four copper–graphite composites containing different graphite content were prepare dusing rectangular tool through repeating the process passes. Friction and wear performance of the composites were also studied using a pin-on-disctribometer. It was indicated that the friction coefficients of composites were lower than pure annealed copper and decreased with increase in graphite content.

I. Galvao, A. Loureiro and D. M. Rodrigues [6] In their work, 1 and 3 mm-thick copper-DHP plates were processed with the aim of simulating surface (SFSP) and bulk (VFSP) processing. The influence of the processing conditions on the microstructure and mechanical properties of the processed materials was analyzed. It was found that the tool geometry, which has a close relation with the plastic deformation and dynamic recrystallization kinetics inside the stirred volume, the processing parameters and the heat exchange conditions, which determine the extent of dynamic recrystallization and annealing phenomenon, are determinant in FSP. Kudzanayi Chiteka [8] studied making a choice in selection of friction stir welding / processing (FSW/P) tool material which has become an important task in determining the quality of the weld produced. The tool material selection depends on the operational characteristics such as temperature, wear resistance and fracture toughness that determine the type of materials to be joined. Soft materials can be easily welded using tool steels while harder materials need harder tool materials such as carbide based materials and polycrystalline cubic boron nitride (PCBN).K. Surekha, A. Els-Botes [9] the objective of their study was to obtain a high strength, high conductivity copper by friction stir processing. Three millimeter thick pure copper plate was friction stir processed to a depth of 2.8 mm at low-heat input conditions by varying the travel speed from 50 to 250 mm/min at a constant rotation speed(300 rpm) to obtain fine grains.

Mohsen Barmouz, et al [11]studied multi-pass friction stir processing (MFSP) which was used for improvement of micro structural and mechanical properties of in situ Cu/SiC composites. Field emission scanning electron microscopy and optical microscopy images indicate that multi-pass FSP notably enhances the separation and dispersion of SiC particles and also reduces the grain size in the composite matrix, SiC particles size and porosity contents. Mohsen Barmouz, et al [12] their study was to produce copper reinforced metal matrix composite (MMC)layers using micron sized SiC particles via friction stir processing (FSP) in order to enhance surface mechanical properties. Microstructural evaluation using optical microscopy (OM) and scanning electron microscopy (SEM) indicated that an increase in traverse speed and a decrease in rotational speed cause a reduction in the grain size of stir zone (SZ) for the specimens friction stir processed without SiC particles. With the aim of determining the optimum processing parameters, the effect of traverse speed as the main processing variable on microstructure and micro hardness of MMC layers was investigated. P. Xue, et al[13]worked with large-area bulk ultrafine grained (UFG) pure Cu which was successfully prepared by multiple-pass overlapping friction stir processing (FSP) under additional rapid cooling. It was observed that overlapping FSP does not exert a significant effect on the microstructure and mechanical properties of the FSP UFG Cu. Similar average grain size was achieved in the transitional zone (TZ) of the multiple-pass FSP sample compared to that in the nugget zone of the single-pass FSP sample, and the TZ exhibited a strong {111} (112) type A fiber shear texture. Q. Zhang, et al [14] investigated situ Al3Ti/Al–5.5Cu composites fabricated by powder metallurgy and subsequent forging which were subjected to multiple pass friction stir processing (FSP) with and without active cooling. Valentin N. Popov [26] studied Carbon nanotubes: properties and application. V Jegannathan et al. [25] investigated the parameters affecting the friction stir processed copper and enhancement of the mechanical properties of the composite material. The results showed that the grain size of fabricated composite reduce, also it is indicated that in comparison to base metal (copper) micro hardness of friction stir processed composites in stir zone (SZ) increase significantly. The results obtained also indicated that the selected FSP parameters significantly influence the area of surface composite by the distribution of material particles. Higher tool rotational speed and lower processing speed produce an excellent distribution of material particles and higher area of surface composite due to higher frictional heat, increased stirring and material transportation. Ranganath M S et al [21] conducted the experimental studies for assessing the tribological performance of aluminium at the sliding contacts with mild steel plate, using a pin-on-disk tribometer as per ASTM-G 99. The study has been done in order to explore the friction and wear behaviors at the interface of tribo-pair. The tribological properties as coefficient of friction and specific wear rate of aluminium 6061are investigated. The Tribological tests are carried out at 500, 1000, 1500 rpm for 1000 meters in dry condition based on Response Surface Methodology. Track diameter, rotating speed and normal load are considered as the design parameters. Using central composite design, the problem is converted into single response optimization problem and the optimum combination of design parameters are found as 50mm track diameter, 500 rpm of rotating speed and 0.5 kg of normal load. The ANOVA result shows that the rotational speed is the most significant factor, followed by load and Track diameter for co-efficient of friction. Whereas the Track diameter is the most significant factor, followed by rotating

speed and normal load for specific wear rate. Finally, microscopic images are investigated to identify the wear mechanism.

R. Sathiskumar, et al [18] applied Friction stir processing (FSP) to fabricate boron carbide (B4C) particulate reinforced copper surface composites. The effect of FSP parameters such as tool rotational speed, processing speed and groove width on microstructure and micro hardness was investigated. A groove was contrived on the 6mm thick copperplates and packed with B4C particles. FSP was carried out using five various tool rotational speeds, processing speeds and groove widths. R. Sathis kumar, et al [19] applied the friction stir processing technique to fabricate boron carbide particulate reinforced copper surface composites and investigate the effect of B4C particles and its volume fraction on microstructure and sliding wear behave or of the same. A groove was prepared on 6 mm thick copper plates and packed with B4C particles. The dimensions of the groove was varied to result in five different volume fractions of B4C particles (0, 6, 12, 18 and 24 vol. %). R. Sathis kumar, et al [20] In this work FSP technique was applied to prepare copper surface composites reinforced with variety of ceramic particles such as SiC, TiC, B4C, WC and Al2O3. Empirical relationships were developed to predict the effect of FSP parameters on the properties of copper surface composites such as the area of the surface composite, micro hardness and wear rate. The FSP parameters which influence the properties of surface composite are shown in Fig. 1. Salar Salahi and Vahid Rezazadeh [22] studied Fracture mechanism in metals using Friction Stir Processing (FSP) which was a challenging investigation and made by means of a rotating tool inserted in a work piece providing heat transfer and plastic deformation. In their paper, improving ductility during FSP was determined as a purpose and the microstructure and fracture mechanism of samples were investigated during Friction Stir Processing (FSP) of pure copper.

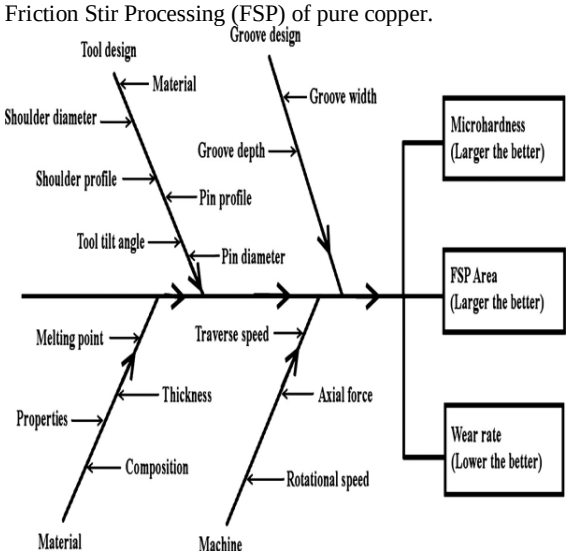


Fig: 1. FSP parameters influencing the properties of surface composite. [20]

3. Experimental Setup

The machine used for Friction stir processing was a special FSW machine is shown below.



Fig: 2. Friction stir welding / processing machine (Central workshop, Delhi Technological University)



Fig: 3. Tools used during Friction stir Processing

The specimen size of the copper plate that is used for processing 200 mm x 74 mm x 5 mm. One groove of 1mm width and 2 mm deep was made on the 99.99% pure Copper Plates using horizontal milling machine with a 1mm (width) saw cutter was cut in the middle of the specimen plate for processing. The tool Material used is H13 steel with shoulder diameter 15mm, threaded pin diameter 8 mm, pin length 2.5 mm with tool rotational speed 960 rpm, tool angle 2^0 and table traverse speed 25 m / min. Initially without carbon nanotubes one specimen plate was processed with single pass. Then second specimen plate was processed after filling carbon nanotubes in the groove cut (single pass). Then third specimen plate was processed after filling carbon nanotubes in the groove cut (double pass). Finally the fourth specimen plate was processed after filling carbon nanotubes in the groove cut (triple pass). The processed pieces were taken for the following tests: Tensile Strength Test, Brinell hardness test, Microstructure test.

4. Tensile Strength Test

After the FSP the specimen for tensile testing were cut from the job. The specimens are as shown in figure. The specification of specimen is as following:

- Length of specimen- 101.6 mm

- Gauge Length –25.4 mm
- Gauge width- 6.00mm
- Gauge thickness- 5.00mm

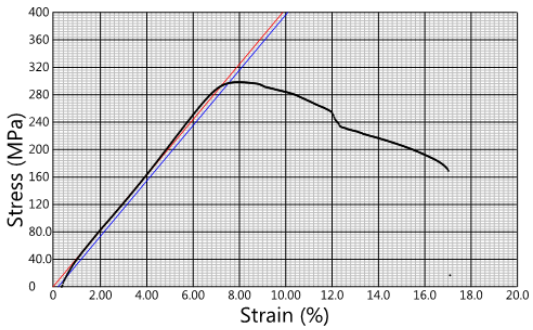
The specimens were chosen marked with marker on their ends. Care was taken to ensure that the specimens did not have any notching or cracks from manufacturing or any surface defects that would adversely affect the tensile tests. Before loading the specimens in the Instron machine, the computer system connected to the machine was given inputs such as gauge length and width of the specimen. The computer system was then prepared to record data and output necessary load-deflection graphs. The specimens were loaded into the Instron machine, and a tensile test was performed. The data was recorded electronically in text files and the load-deflection curve was shown on the computer screen as a visual representation. The average of different values of 3 specimens each from 2 jobs and parallel material was taken as the final values. The stress strain graphs and load elongations graph are shown for the ultimate strength of the specimen. The ultimate tensile strength of the processed material comes out to be lesser than the parent material.

Table: 1a. Specimens Details and Output Results during Tensile Tests

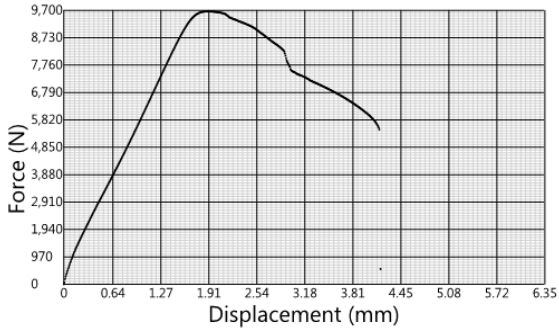
Out Put (Generic metals tensile from position	**Without processing**	**Single Pass Without CNT**
Width (mm)	6.00	6.00
Thickness (mm)	5.40	5.30
Gauge Length (Initial) (mm)	25.0	25.0
Gauge Length (Final) (mm)	28.7	31.8
Area (mm^2)	32.4	30.8
Ultimate Force (N)	9670	5650
Ultimate Stress (MPa)	298	183
Offset @ 0.2% (N)	9620	1240
Offset @ 0.2% (MPa)	297	40.2
TE (Auto) (%)	17.1	26.3

Table: 1b. Specimens Details and Output Results during Tensile Tests

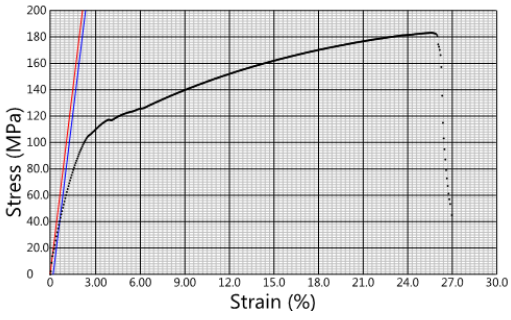
Out Put (Generic metals tensile from position	**Single Pass With CNT**	**Two Pass With CNT**	**Three Pass With CNT**
Width (mm)	6.00	6.00	6.00
Thickness (mm)	5.14	5.05	5.00
Gauge Length (Initial) (mm)	25.0	25.0	25.0
Gauge Length (Final) (mm)	26.8	27.4	30.4
Area (mm^2)	30.8	30.3	30.0
Ultimate Force (N)	2800	4860	6380
Ultimate Stress (MPa)	90.7	160	213
Offset @ 0.2% (N)	2550	2120	4540
Offset @ 0.2% (MPa)	82.7	70.0	151
TE (Auto) (%)	7.28	9.72	22.2



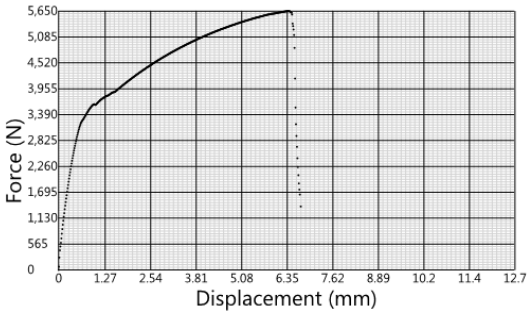
(a) Stress-strain curve for specimen without processing



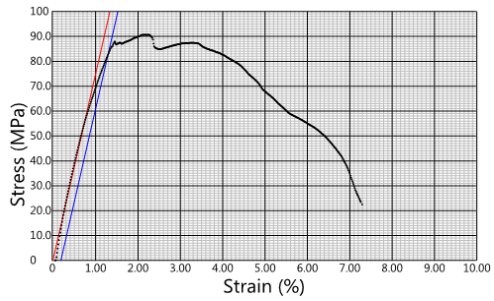
(b) Force-displacement curve for specimen without processing



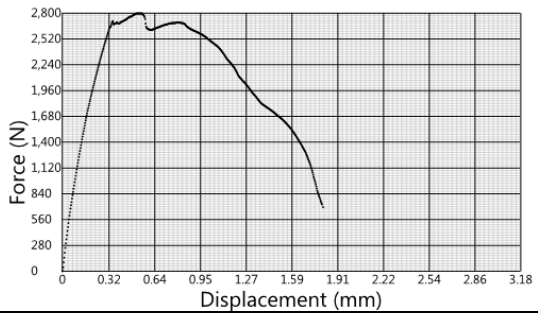
(c) Stress-strain curve for specimen without cnt single pass processing



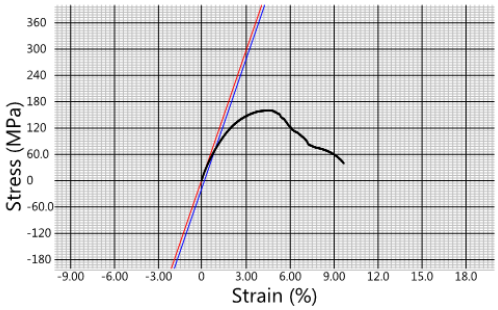
(d) Force-displacement curve for specimen without cnt single pass processing



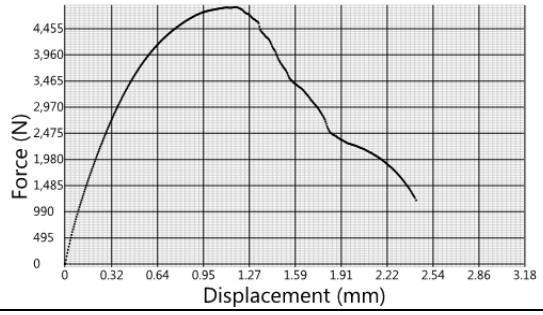
(e) Stress-strain curve for specimen with cnt single pass processing



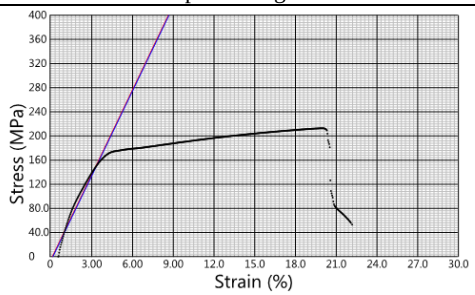
(f) Force-displacement curve for specimen with cnt single pass processing



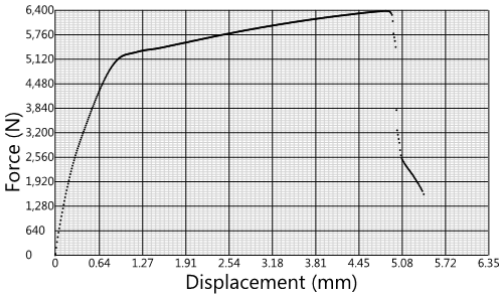
(g) Stress-strain curve for specimen with cnt double pass processing



(h) Force-displacement curve for specimen with cnt double pass processing



(g) Stress-strain curve for specimen with cnt triple pass processing



(h) Force-displacement curve for specimen with cnt triple pass processing

Fig: 4. output results during tensile tests after Friction stir processing

5. Brinell Hardness Test

The Brinell scale characterizes the indentation hardness of materials through the scale of penetration of an indenter, loaded on a material test-piece. It is one of several definitions of hardness in materials science. The brinell hardness was conducted on all specimens .As we can see the brinell hardness of the processed specimen increases as the number of passes are increased due to the more compact microstructure. The brinell hardness number comes out to be highest for third pass in comparing with single and double passes.

Table: 2. Specimens Details and Output Results during Tensile Tests

Expts	Description	Depth of indendation	Brinell hardness BHN
1	Without carbon nanotube and without processing	2.8 mm	11.14
2	single pass processing Without carbon nanotube	3.09	8.93
3	Single pass With carbon nanotube	3.47	6.823
4	Double passes With carbon nanotube	3.40	7.162
5	Triple passes With carbon nanotube	3.00	9.55

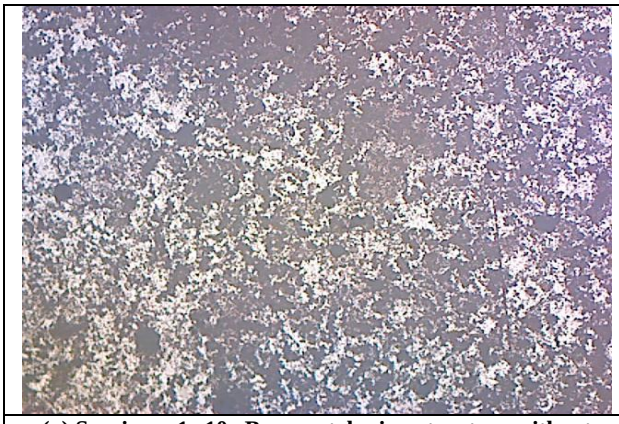
6. Microstructure Ananlyis

The optical microscope is the principal tool used to characterize the internal grain structure of steel. Traditionally, the structure revealed by the microscope is called the microstructure. The mechanical properties of given steel are strongly influenced by its microstructure. An optical microscope uses reflected light to generate an image.

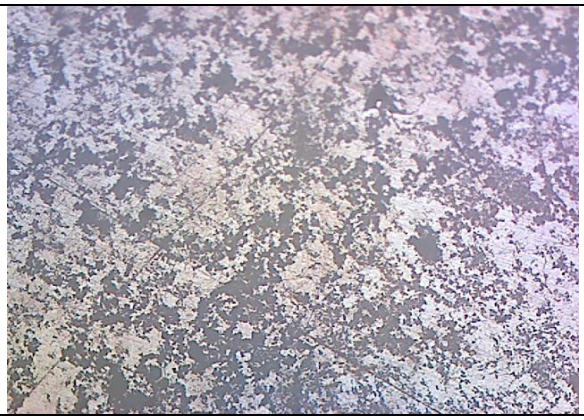
A beam of light is directed down onto the surface and the image is generated either on film or the eye by light reflected along the same direction.This is performed on specimens either cut to size or mounted in a resin mold. The samples are polished to a fine finish, normally one micron diamond paste, and usually etched in an appropriate chemical solution prior to examination on a metallurgical microscope. For microstructure test we had five specimens. There are two examination methods in metallography:

(i) Macroscopy
(ii) Microscopy

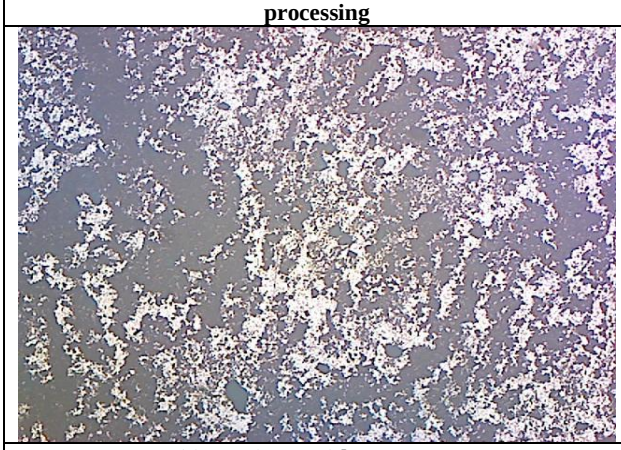
In macroscopy the examination of the structural characteristics or chemical characteristics of a metal or an alloy is done by the unaided eye or with the aid of a low-power microscope or binocular, usually under l0x. In microscopy similar examination is done with the prepared metal specimens, employing magnifications with the optical microscope of from 100 x to as high as 2000x. Figure 5 represents the microscopic results of all 5 specimens after processing



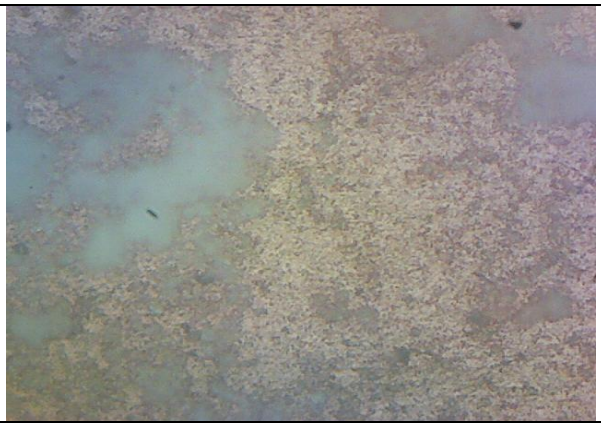
(a) Specimen 1--10x Base metal microstructure without processing



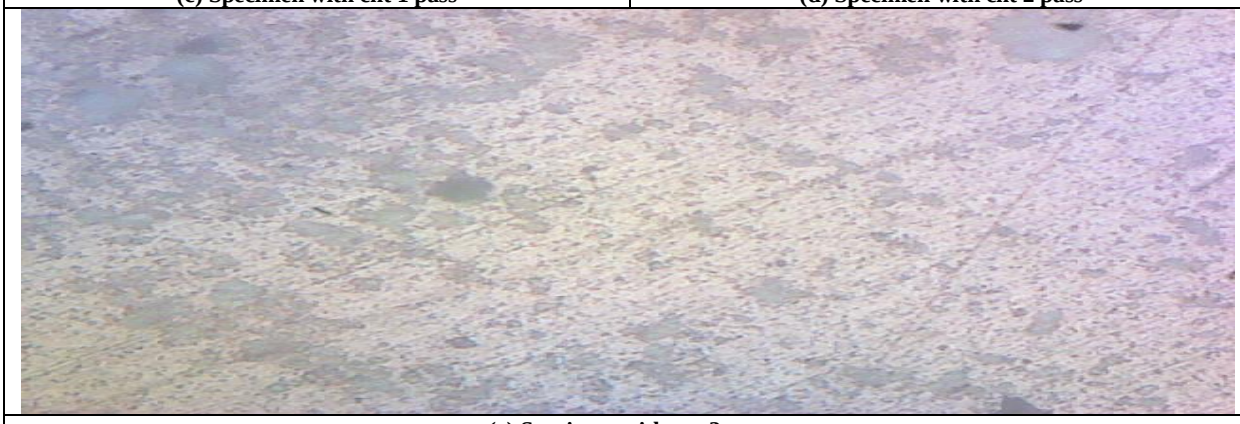
(b) Specimen 1--20x Specimen without cnt with 1 pass



(c) Specimen with cnt 1 pass



(d) Specimen with cnt 2 pass



(e) Specimen with cnt 3 pass

Fig: 5. a-e. Microscopic Results of all Specimens After Friction Stir Processing

7. Conclusions

The effect of friction stir processing (FSP) parameters on copper material with carbon nano tubes as composite material has been studied. Tensile strength, microstructure and micro hardness tests were conducted on the specimens prepared at the central workshop of Delhi Technological University, India. The following conclusions have been made from present investigation.

- Fabrication of copper–carbon nano tube composites by friction stir processing (FSP) is possible.
- The microstructure for different specimens showed that after single, double and triple passes the microstructure gets more compact and observed no defects after processing.
- The brinell hardness number comes out to be highest for third pass in comparing with single and double passes.

References

[1] B. A. Khiyavi, A. J. Aghchai, M. Arbabtafti, M. K. B. Givi, J. Jafari, Effect of friction stir processing on mechanical properties of surface composite of Cu reinforced with Cr particles, Advanced Materials Research, 829, 2014, 851-856

[2] B. C. Liechty, B. W. Webb, Flow field characterization of friction stir processing using a particle-grid method, journal of materials processing technology, 208, 2008, 431–443

[3] G. Chai, Y. Sun, Jianren 'Jenny' Sun and Quanfang Chen, Mechanical properties of carbon nanotube–copper nanocomposites, IOP Science, Journal of Micromechanics and Microengineering Volume 18 Number 3, Guangyu Chai et al 2008 J. Micromech. Microeng. 18 035013

[4] H.R. Akramifard, M. Shamanian, M. Sabbaghian, M. Esmailzadeh, H. R. Akramifard, M. Shamanian, M. Sabbaghian, M. Esmailzadeh, Microstructure and mechanical properties of Cu/SiC metal matrix composite fabricated via friction stir processingMaterials and Design, 2013

[5] H. Sarmadi, A. H. Kokabi, S. M. SeyedReihani, Friction and wear performance of copper–graphite surface composites fabricated by friction stir processing (FSP) , Wear 304, 2013, 1–12

[6] I. Galvao, A. Loureiro, D. M. Rodrigues, Influence of process parameters on the mechanical enhancement of copper-DHP by FSP, Advanced Materials Research, 445, 2012, 631-636

[7] Jian-Qing Su, TW Nelson, TR McNelley, R S Mishra, Development of nanocrystalline structure in Cu during friction stir processing (FSP), Materials Science and Engineering: A, Elsevier, 2011/6/25, 5458-5464

[8] K. Chiteka, Friction Stir Welding/Processing Tool Materials and Selection, International Journal of Engineering Research & Technology, 2(11), 2013, 8-18

[9] K. Surekha , A. Els-Botes, Development of high strength, high conductivity copper by friction stir processing, Materials and Design, 32, 2011, 911–916

[10] L. B Johannes, L. L Yowell, E. Sosa, S. Arepalli, R. S Mishra, Survivability of single-walled carbon nanotubes during friction stir processing, IOP Science, Nanotechnology, Lucie B Johannes et Nanotechnology 17 3081, 2006

[11] M. Barmouz, M. K. B. Givi, Fabrication of in situ Cu/SiC composites using multi-pass friction stir processing: Evaluation of microstructural, porosity, mechanical and electrical behaviour, Composites: Part A 42, 2011, 1445–1453

[12] M. Barmouz, M. K. B. Givi, J. Seyfi, On the role of processing parameters in producing Cu/SiC metal matrix composites via friction stir processing: Investigating microstructure, icrohardness, wear and tensile behaviour, Materials Characterization 62, 2011, 108 – 117

[13] P. Xue, B. L. Xiao, Z. Y. Ma, Achieving Large-area Bulk Ultrafine Grained Cu via Submerged Multiple-pass Friction Stir Processing, J. Mater. Sci. Technol., 29(12), 2013, 1111-1115

[14] Q. Zhang, B. L. Xiao, P. Xue, Z. Y. Ma, Microstructural evolution and mechanical properties of ultrafine grained Al3Ti/Al–5.5Cu composites produced via hot pressing and subsequent friction stir processing, Materials Chemistry and Physics 134, 2012, 294– 301

[15] Rajiv S. Mishra, Murray W. Mahoney, Friction Stir Welding and Processing, ASM International, 2007

[16] Rajiv S Mishra, Bulk nanomaterials from friction stir processing: Features andproperties, Book: Bulk Nanostructured Materials, Wiley-VCH Verlag GmbH & Co KGa A Weinheim, 255-272

[17] R S Mishra, H Jones, GW Greenwood, An empirical correlation for the grain-boundary diffusion of impurities in copper, Journal of materials science letters, 1988/7/1, 728-730

[18] R Sathiskumar, N Murugan, I. Dinaharan, S. J. Vijay, Role of friction stir processing parameters on microstructure and microhardness of boron carbide particulate reinforced copper surface composites, Indian Academy of Sciences, Sadhana, 38(6), 2013, 1433–1450

[19] R. Sathiskumar, N. Murugan, I. Dinaharan, S. J. Vijay, Characterization of boron carbide particulate reinforced in situ copper surface composites synthesized using friction stir processing, Materials Characterization, 84, 2013, 16 – 27

[20] R. Sathiskumar, N. Murugan, I. Dinaharan, S. J. Vijay, Prediction of mechanical and wear properties of copper surface composites fabricated using friction stir processing, Materials and Design 55, 2014, 224–234

[21] Ranganath. M. S, R. C. Singh , Rajiv Chaudhary, R. K. Pandey, Experimental Investigation of Friction and Wear Behavior at the Interface of Aluminium and Mild Steel, International Journal of Advance Research and Innovation, 2(4), 2014, 775-780

[22] S. Salahi, V. Rezazadeh, Fracture Mechanism in Friction Stir Processed Annealed Pure Copper

Samples, World Applied Sciences Journal, 23(12), 2013, 54-58

[23] S. F. Miller, New friction stir techniques for dissimilar materials processing, Manufacturing Letters 1, 2013, 21–24

[24] V. Jeganathan Arulmoni, Ranganath M S , R S Mishra, Effect of Process Parameters on Friction Stir Processed Copper and Enhancement of Mechanical Properties of the Composite Material: A Review on Green Process Technology, International Research Journal Of Sustainable Science & Engineering, IRJSSE, 2(4), 2014

[25] V. Jeganathan Arulmoni, R. S. Mishra, Ranganath M. S, Experimental Investigations on Friction Stir Processed Copper and Enhancement of Mechanical Properties of the Composite Material, International Journal of Advance Research and Innovation, 2(3), 2014, 557-563

[26] Valentin N. Popov, Carbon nanotubes: properties and application, Materials Science and Engineering, R 43, 2004, 61–102

[27] Yan-Hui Li, William Housten, Yimin Zhao and Yan Qiu Zhu, Cu/single-walled carbon nanotube laminate composites fabricated by cold rolling and annealing, IOP Science, Nanotechnology, Yan-HuiLi et al 2007 Nanotechnology 18 205607, 2007, IOP Publishing Ltd

[28] Yong X. Gan , Daniel Solomon 2 and Michael Reinbolt, Friction Stir Processing of Particle Reinforced Composite Materials, Materials 2010, 3, 329-350

[29] Z. Y. MA, Friction Stir Processing Technology: A Review, The Minerals, Metals & Materials Society and ASM International, 2008

Int J Adv Manuf Technol
DOI 10.1007/s00170-015-6996-3

ORIGINAL ARTICLE

Experimental investigation into tool aging effect in ball end magnetorheological finishing

Mahendra Singh Niranjan[1] · **Sunil Jha**[2]

Received: 17 December 2013 / Accepted: 5 March 2015
© Springer-Verlag London 2015

Abstract Magnetorheological (MR) finishing is a smart finishing processes applied to a variety of applications. In the present work, an attempt has been made to improve the percentage reduction in surface roughness with sintered magnetic abrasives based MRP fluid on BEMRF tool. Sintered magnetic abrasives were developed by mixing of 20 vol% carbonyl iron powder (CIP) CS grade and 25 vol% SiC in ball mill then pellets of 5 g each have been made at 8 ton pressure. These pellets were allowed to sinter in tubular furnace at 1200 °C in the controlled atmosphere of argon. The sintered pellets were crushed in a ball mill to obtain sintered magnetic abrasives. The morphology and magnetizability of sintered magnetic abrasives have been studied with scanning electron microscope (SEM) and vibration sample magnetometer (VSM). MRP fluid was synthesized with 45 vol% sintered magnetic abrasives and 55 vol% base fluid. Experiments were conducted on mild steel work-piece for 30 min using BEMRF tool at given machining conditions. Percentage reduction in surface roughness ($\%\Delta R_a$) was calculated and compared with $\%\Delta R_a$ obtained by finishing mild steel work-piece with unbonded magnetic abrasives based MRP fluid and found better results. The effect of rotation per minute (RPM) on $\%\Delta R_a$ has been studied and found optimum by conducting the experiments at higher values of RPM keeping other machining conditions same. The experimental study shows that the sintered magnetic abrasives based MRP fluid in MR finishing may be used at higher rotational speeds without tool aging effect and impart a higher finishing force on the work-piece surface, improves finishing capability of the process.

Keywords Magnetorheological · MRP fluid · Nanofinishing · Sintering · Magnetic abrasives

Nomenclature

R_a	Centre line average surface roughness (μm)
R_q	Root mean square (rms) surface roughness (μm)
R_z	Peak to valley height surface roughness (μm)
MR	Magnetorheological
BEMRF	Ball end magnetorheological finishing
RPM	Rotation per minute
CIP	Carbonyl iron powder
$\%\Delta R_a$	Percentage reduction in surface roughness
VSM	Vibration sample magnetometer
SEM	Scanning electron microscope
SiC	Silicon carbide abrasives

1 Introduction

Magnetorheological (MR) finishing process is based on MR polishing fluid which is composed of micron size magnetic carbonyl iron powder (CIP) and nonmagnetic polishing abrasives in carrier fluid. Magnetorheological (MR) fluids are smart fluids which behave as Newtonian fluids in the absence of magnetic field. The magnetorheology of MR fluid changes reversibly from fluid to semi-solid in a few milliseconds on the application of magnetic field [1]. With this behaviour, MR fluid based finishing process has been applied to finish a variety of materials such as glass etc. to hard metals. In this process, the surfaces are selectively finished by changing the MR fluid's yield shear stress with externally applied magnetic

✉ Sunil Jha
suniljha@mech.iitd.ac.in

[1] Delhi Technological University, Bawana Road, New Delhi, Delhi, India

[2] I.I.T. Delhi, New Delhi, Delhi, India

Published online: 01 May 2015

field and nanolevel surface finish was achieved on optical lenses [2]. The process has been used to obtain the curved surfaces of silicon-based micro-structure and finishing three dimensional components. Finite element method (FEM) was employed to investigate the magnetic field effect near the tool assembly on the finishing surface profile [3, 4]. Optical aspheric components were polished with MR fluid assistive finishing process. Experiments were conducted on K9 glass parabolic lens to see the parametric effect on material removal and performance [5]. The experimental study was carried out to see the parametric effect on surface finish and material sheared in MR finishing of silicon blank [6]. For in-depth understanding of the mechanism of material sheared in MR finishing, a mathematical model for the forces acting on work-piece has been proposed. Two different theories were used, and validation of model has been done with experimental results. Force measurement on free form surface was carried out experimentally, and parametric effect on tangential, normal, and axial forces was studied [7, 8]. A novel ball end magnetorheological finishing (BEMRF) process was developed to finish flat and three dimensional surfaces of ferromagnetic and diamagnetic work-piece. There is no restriction on relative movement between finishing medium and the work-piece surface [9]. A three-dimensional mild steel work-piece surface was finished with the BEMRF process, and performance of the process was evaluated by studying the effect of number of finish passes on final surface roughness. The analysis of process parameters on BEMRF process has been carried out, and response on percentage reduction in surface roughness ($\%\Delta R_a$) has been seen. It was observed that $\%\Delta R_a$ continuously reduces with increasing tool rotational speed because the polishing abrasives embedded by surrounding CIPs are prone to escape from their position at high tool rotational speed due to centrifugal force [10, 11]. For in-depth understanding, the mechanism of material sheared and surface finish on the ferromagnetic work-piece, a mathematical model for magnetic field induced normal force has been developed in BEMRF process [12]. A new kind of sintered iron-CNT compound abrasives was developed for MR finishing of hard work-piece surface. The parametric study has been done to improve the rate of material removal, and tool aging effect has been minimized [13]. CIP and silicon carbide (SiC) based sintered magnetic abrasives has been developed in tubular furnace at 1000 °C in an inert atmosphere of argon using pressureless solid phase sintering method. The bulk density of sintered magnetic abrasives has been obtained by Archimedes method. Then, MR polishing fluid was prepared and experiment was conducted with optimum machining conditions on BEMRF tool [14]. For in-depth understanding, the CIP and SiC based sintered magnetic abrasives has been characterized to see its magnetic properties, morphology, elemental composition and different phases. After characterization, MR polishing fluid was synthesized and experiments were conducted on mild steel work-piece with BEMRF tool to study $\%\Delta R_a$ and tool aging effect [15]. In the existing unbonded magnetic abrasives based MR polishing fluid finishing processes, it has been reported that the micro structural cavities are formed due to association of large CI particles and abrasives. These micro-cavities were filled with the use of bidisperse MR polishing fluids which were prepared with different grades of CIP (CS and HS grades obtained from BASF Germany) with different volume fraction combinations within 20 vol% total magnetic content and rest is SiC abrasives and base fluid. Then, flow behaviour of prepared MR polishing fluid has been studied using magnetorheometer at different magnetic field strength and steady-state rheograms have been drawn. The result shows that the micro-structure cavities were filled with HS grade of CIP and enhanced the yield shear stress and viscosity of MR polishing fluids [16].

Table 1 MR polishing fluid compositions

Sample number	MR polishing fluid compositions
Sample 1 (unbonded)	20 vol% CIP CS grade, 25 vol% SiC abrasives, and 55 vol% base fluid
Sample 2 (sintered)	45 vol% sintered magnetic abrasives and 55 vol% base fluid

From the literature review, it is revealed that the polishing abrasives in existing MR fluid based finishing processes, are prone to escape from the surface of MR polishing fluid under high tool rotational speed due to centrifugal force acting on it and result in reduced rate of material removal, referred as tool aging effect. Hence, the existing MR finishing is performed at low tool rotational speed due to tool aging effect and reduce productivity.

To further improve the $\%\Delta R_a$ as well as overall reliability in MR finishing, an attempt has been made to use sintered magnetic abrasives based MR polishing fluid with BEMRF process. In the present work, the experiments were conducted on mild steel work-piece surface with sintered magnetic abrasives based MR polishing fluid using BEMRF process at 332 rotation per minute (RPM) and other machining parameters given in Table 2. Then $\%\Delta R_a$ was calculated and compared with $\%\Delta R_a$ obtained by finishing the mild steel surface with unbonded magnetic abrasives based MR polishing fluid at same machining conditions. Unbonded magnetic abrasives

Table 2 Experimental parameters and other conditions

Parameters	Conditions
Tool rotational speed (RPM)	332, 400, 500, 600, 668
Working gap (mm)	0.75
DC power supply	4 A
Total finishing time (minute)	30
Abrasive	Silicon carbide, 800 mesh size
Work-piece material	Mild steel

Table 3 Surface roughness before and after MR finishing

MRP fluid sample number	Initial roughness (μm)	Final roughness (μm)	%Reduction in roughness
Sample 1 (unbonded)	R_a 0.2163	R_a 0.1755	19 %
	R_q 0.2753	R_q 0.2213	20 %
	R_z 1.7240	R_z 1.5187	15 %
Sample 2 (sintered)	R_a 0.2141	R_a 0.1444	32.55 %
	R_q 0.2659	R_q 0.1885	29.11 %
	R_z 1.3155	R_z 0.9560	27.32 %

are uniform mixture of 20 vol% CIP of CS grade (BASF) and 25 vol% SiC abrasives. Effect of tool rotation speed was studied with sintered magnetic abrasives based MR polishing fluid to see tool aging effect and $\%\Delta R_a$ on BEMRF process.

2 Development of sintered magnetic abrasives

Sintered magnetic abrasives were developed by uniformly mixing of 20 vol% CIP CS grade and 25 vol% SiC abrasives in ball mill; then, pellets of 5 g each have been made at 8 ton pressure and hold at this pressure for 20 min to get uniform compaction. These pellets were allowed to sinter in a tubular furnace in the controlled atmosphere of argon with temperature increment at 4 °C per minute, reach a temperature of 800 °C, and hold at this temperature for 1 h for getting complete degasification. After degasification, the temperature was further increased with the same rate up to 1200 °C and is held at this temperature for 3 h for getting complete phase transformation. At this stage, faster rate of solid state surface diffusion of CIP and SiC grains takes place. CIP reacts with SiC and forms iron silicides at the interfaces [17]. After sintering, the furnace cooling was done in argon atmosphere till the room temperature of 25 °C was achieved. The sintered pellets were crushed in a ball mill to obtain sintered magnetic abrasives. Hence, sintered magnetic abrasives composed of ferromagnetic and abrasive phase combined firmly were obtained.

3 Characterization

Morphology of unbonded and sintered magnetic abrasives was studied by scanning electron microscope (SEM). VSM has been used to study the magnetizability of unbonded and sintered magnetic abrasives in powder form at 25 °C. The bulk density of sintered abrasives was determined by Archimedes method. Five samples were tested and average bulk density was found as 5.5 g/cm^3.

4 Preparation of samples

MR polishing fluid samples were synthesized with unbonded magnetic abrasives of 20 vol% CIP CS grade, 25 vol% SiC abrasives, and 55 vol% base fluid (sample 1), and another sample consists of 45 vol% sintered magnetic abrasives and 55 vol% viscoplastic base medium (20 weight% AP3 grease and 80 weight% paraffin liquid heavy) (sample 2) as given in Table 1.

5 Mechanism of material removal

The synthesized MR polishing fluid was passed by peristaltic pump and enters axially under pressure from upper end of BEMRF tool, reaches at the tool tip. Once the MR polishing fluid reached at MR finishing tool tip, the sintered magnetic abrasives together with crushed CIPs and/or surrounding

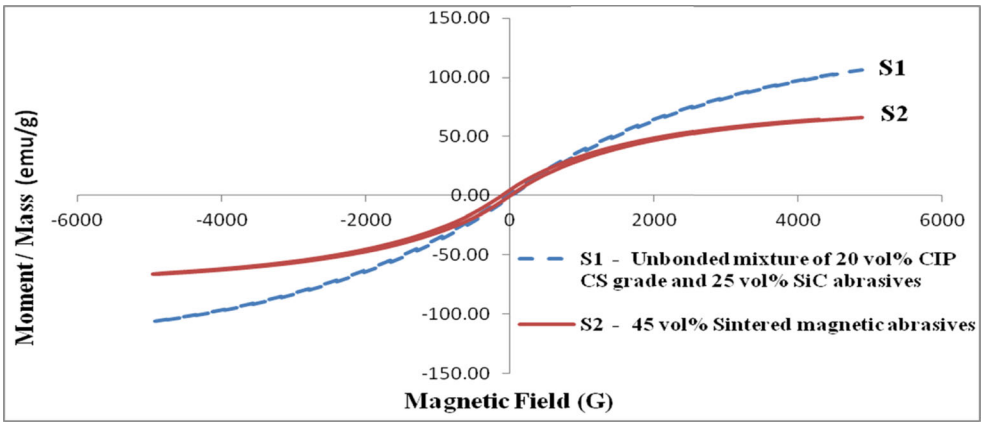


Fig. 1 Vibration sample magnetometery (VSM) data for unbonded and sintered magnetic abrasives

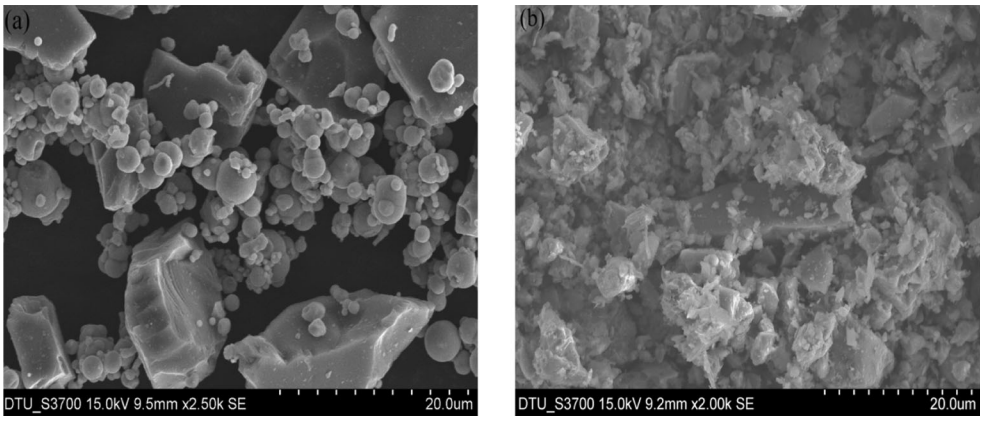


Fig. 2 Scanning electron microscopy of unbonded and sintered magnetic abrasives

sintered magnetic abrasives in MR polishing fluid are aligned in magnetic field direction and form a strong chain like columnar structure under external magnetic field. Thus, MR polishing fluid stiffened in the hemisphere shape as shown in Fig. 3a. The movement of stiffened MR polishing fluid ball end on the work-piece surface is enabled by motion controller integrated with the computer. During the MR finishing process, the forces are acting directly on the sintered magnetic abrasives as shown in Fig. 3b. The magnetic force between sintered magnetic abrasives provides bonding strength to it. When stiffened, MR polishing fluid ball end has relative motion with work-piece surface during MR finishing tool rotation, it shears off the peaks from the work-piece surface. The amount of material sheared from peaks of the work-piece surface by sintered magnetic abrasive grains depends upon bonding force (F_n) between magnetic abrasive grains provided by magnetic field in ball end MR polishing fluid and shear force (F_t) due to MR finishing tool rotational speed. The normal force F_n provided by magnetic field induced structure is responsible for depth of penetration whereas shear force F_t is responsible for material removal. The removal of material in μ-chips takes place under the influence of resultant finishing force F_F given by [9]

The resistance offered by the work-piece surface for material removal is given by

$$F_{req} = \tau_s . A_p \tag{2}$$

where τ_s is the work-piece shear strength and A_p is the projected area of penetration.

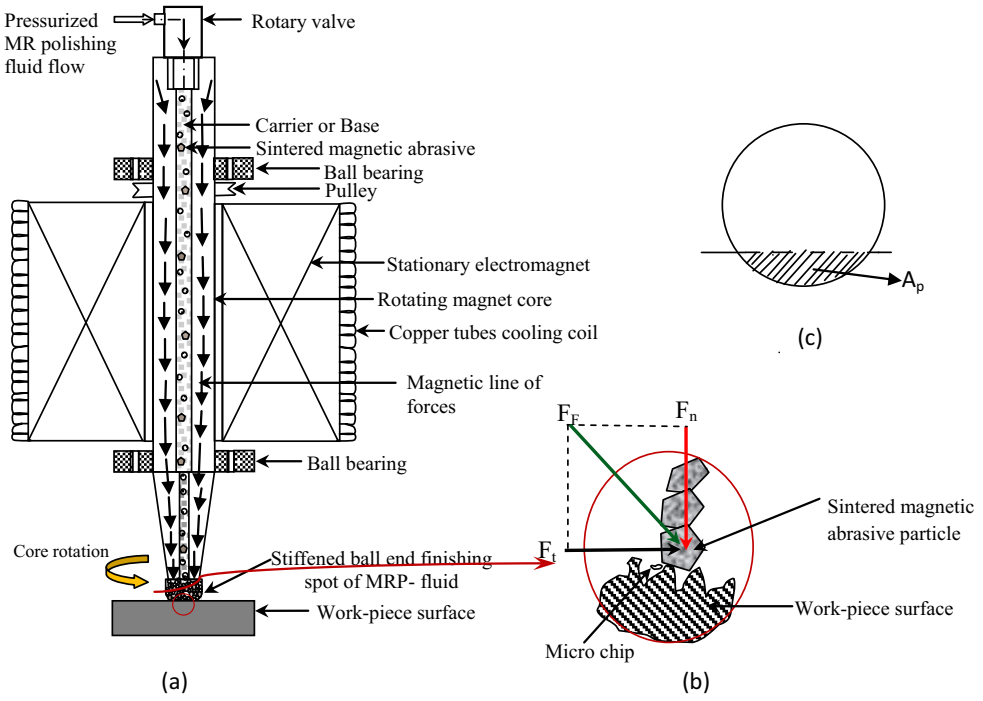


Fig. 3 Schematic diagram of forces acting on sintered magnetic abrasive particle

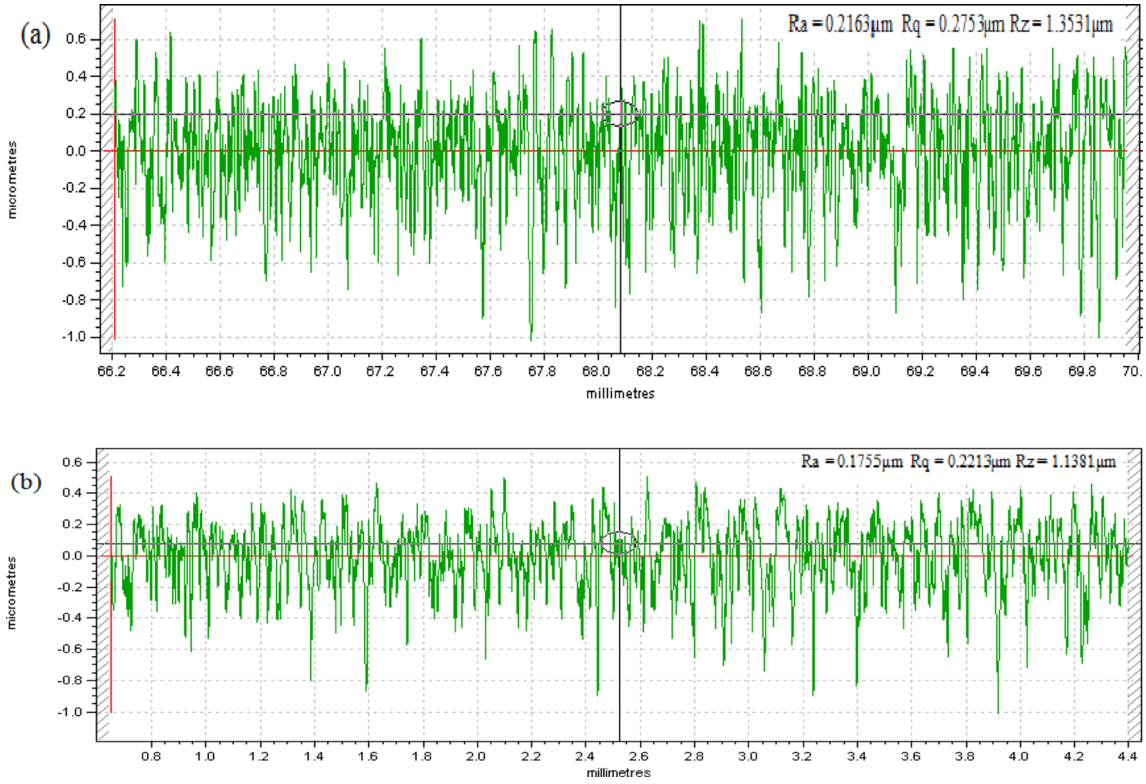


Fig. 4 Surface roughness profile **a** before finish and **b** after finish using MR polishing fluid sample 1

The material is sheared if and only if the resultant finishing force overcomes the resistance offered by work-piece. That is

$$F_F > F_{req} \tag{3}$$

The MR finishing mechanism with sintered magnetic abrasives may be treated as two body wear mechanism where the bonding strength of sintered magnetic abrasives with the neighbouring crushed CIPs and/or sintered magnetic abrasives may be more, as both the sintered magnetic abrasives and surrounding particles are magnetizable. Moreover, sintered magnetic abrasives impart more finishing force on the work-piece surface and shear off peaks of work-piece material more quickly and efficiently.

6 Experimentation

The experiments were conducted on mild steel work-piece for 30 min with unbonded magnetic abrasives based MR polishing fluid (sample 1) and synthesized sintered magnetic abrasives based MR polishing fluid (sample 2) at MR finishing tool rotational speed of 332 RPM, and other

Table 4 Surface roughness before and after MR finishing with sample 2 at different RPM

MR finishing tool rotation speed (RPM)	Initial roughness (µm)	Final roughness (µm)	%Reduction in surface roughness
400	R_a 0.2086	R_a 0.1289	38.21 %
	R_q 0.2595	R_q 0.1637	36.92 %
	R_z 1.2927	R_z 0.8752	32.29 %
500	R_a 0.2241	R_a 0.1132	49.48 %
	R_q 0.2782	R_q 0.1495	46.26 %
	R_z 1.3835	R_z 0.8005	42.14 %
600	R_a 0.2147	R_a 0.0914	57.43 %
	R_q 0.2644	R_q 0.1205	54.43 %
	R_z 1.3110	R_z 0.7243	44.75 %
668	R_a 0.2160	R_a 0.1078	50.09 %
	R_q 0.2654	R_q 0.1438	45.82 %
	R_z 1.3243	R_z 0.7567	42.86 %

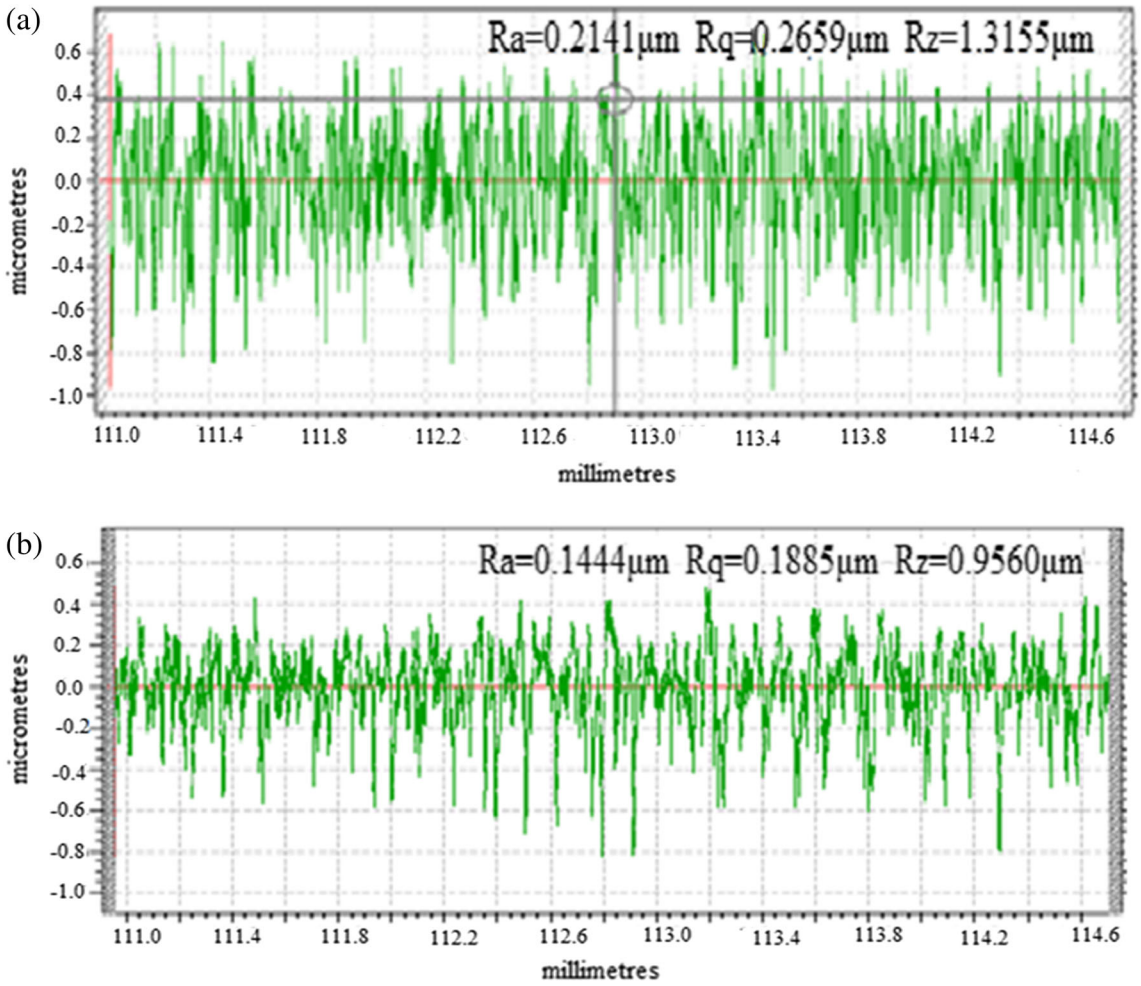


Fig. 5 Surface roughness profile **a** before finish and **b** after finish using MR polishing fluid sample 2

parametric conditions are given in Table 2. The surface roughness before and after experiments was measured by Talysurf at the same point with cutoff evaluation length 0.25 mm and data length 4 mm and given in Table 3. The experiments were further conducted with sintered magnetic abrasives based MR polishing fluid on mild steel work-piece at higher MR finishing tool rotational speed of 400, 500, 600 and 668 RPM keeping other parametric conditions same to see the tool aging effect as well as $\%\Delta R_a$ with tool rotational speed and, hence, optimum MR finishing tool rotational speed has been found.

7 Result and discussion

Figure 1 shows magnetization curves of prepared sintered magnetic abrasives and unbonded magnetic abrasives for same composition by VSM. Maximum magnetization was observed 106.43 emu/g for unbonded magnetic abrasives and 65.8 emu/g for sintered magnetic abrasives at 4893.9 gauss magnetic field. The reduction in magnetization is due to phase change of carbonyl iron from ferromagnetic to paramagnetic (austenite region) on heating at high temperatures then transforming into ferrite and coarse pearlite after furnace cooling [17]. The silicon carbide reacts with carbonyl iron at a faster rate and form iron silicides during sintering at 1200 °C. The morphology of unbounded and sintered magnetic abrasives has been studied by SEM as shown in Fig. 2. The sharp edges of sintered magnetic abrasives are again exposed after crushing the sintered pellets in a ball mill, and shapes and sizes of different magnetic abrasive grains are

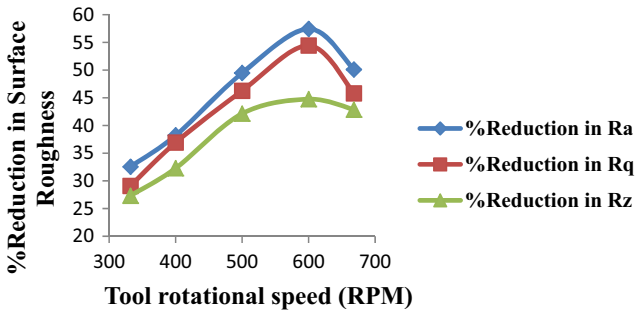


Fig. 6 Effect of tool rotational speed on %Reduction in surface roughness

shown in Fig. 2. During the synthesized MR polishing fluid based finishing process, the shearing of material takes place under the influence of resultant finishing force which is the vector sum of normal and tangential force as shown in Fig. 3b. After the experiments conducted on BEMRF tool, the surface roughness before and after the experiments were measured by Talysurf and shown in Figs. 3 and 4.

Figure 3a, b indicates profile of surface roughness before and after MR finishing for 30 min with unbonded magnetic abrasives based MR polishing fluid (sample 1) at given machining parameters. Then $\%\Delta R_a$ was calculated and found as 19 % in R_a, 20 % in R_q, and 15 % in R_z.

Figure 4a, b indicates the profile of surface roughness before and after MR finishing for a period of 30 min with sintered magnetic abrasives based MR polishing fluid (sample 2) at same machining parameters. Then, $\%\Delta R_a$ was calculated and found as 32.55 % in R_a, 29.11 % in R_q, and 27.32 % in R_z. These results have been compared with $\%\Delta R_a$ obtained by finishing the work-piece surface with unbonded magnetic abrasives based MR polishing fluid. It was observed that the $\%\Delta R_a$ has been found better by finishing work-piece surface with sintered magnetic abrasives based MR polishing fluid (sample 2) than unbonded magnetic abrasives based MR polishing fluid (sample 1). The improvement in shearing action of work-piece material is due to strong bonding between sintered magnetic abrasives and surrounding crushed CIPs and/or sintered magnetic abrasives which impart more finishing force on the work-piece surface.

After comparison, the effect of BEMRF tool rotational speed (RPM) has been studied by conducting the experiments at higher values of RPM keeping other parameters same and $\%\Delta R_a$ has been calculated as shown in Table 4. It was observed that the $\%\Delta R_a$ first increases with RPM and reaches an optimum value, thereafter decreases as shown in Fig. 5. This shows that as the MR finishing tool rotational speed (RPM) increases, shear force F_t acts on abrasive particles increases, which ultimately increases the $\%\Delta R_a$ and reached at optimum level thereafter decreases with further increase of RPM (Fig. 6). This reduction in $\%\Delta R_a$ is due to the occurrence of a tool aging effect at higher tool rotational speed beyond the point of optima. Sintered magnetic abrasives at higher tool rotational speed beyond the optimum level start throw away from their position and reduce the $\%\Delta R_a$ as well as finishing action efficiency. However, the $\%\Delta R_a$ was found continuously decreasing by finishing the work-piece surface with unbonded magnetic abrasives based MR polishing fluid on BEMRF tool [11].

8 Conclusions

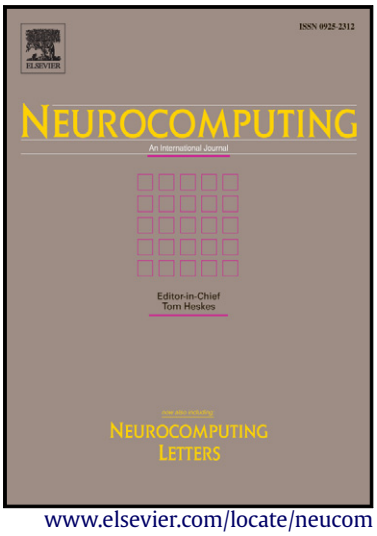
1. Sintered magnetic abrasives were developed at 1200 °C; then morphology, shape, and size have been studied by SEM. It was observed that the CI particles are diffused on the surface of SiC abrasives during sintering. Sharp edges with irregular shapes of sintered magnetic abrasives were found after crushing sintered pellets in a ball mill.
2. Magnetizability has been studied by VSM and observed that maximum magnetization for sintered magnetic abrasives was found less as compare to unbonded magnetic abrasives.
3. Unbonded and sintered magnetic abrasives based MR polishing fluid samples were synthesized.
4. Experiments were conducted on BEMRF tool on mild steel surface with sintered magnetic abrasives and unbonded magnetic abrasives based MR polishing fluid. $\%\Delta R_a$ was calculated after experiments and found better by finishing the work-piece surface with sintered magnetic abrasives based MR polishing fluid as compared to unbonded magnetic abrasives based MR polishing fluid.
5. The effect of tool rotational speed on $\%\Delta R_a$ and tool aging effect has been studied by conducting experiments at higher speed with sintered magnetic abrasives based MR polishing fluid. It was observed that $\%\Delta R_a$ increases with increase of tool rotational speed, then reaches a maximum at 600 RPM, and decreases thereafter due to tool aging effect. Hence, optimum tool rotational speed was found as 600 RPM.
6. Tool aging effect has been minimized with the use of sintered magnetic abrasives based MR polishing fluid.

References

1. Bossis G, Lacis S, Meunier A, Volkova O (2002) Magnetorheological fluids. J Magn Magn Mater 252:224–228
2. Kordonski W-I, Jacobs S-D (1996) Magnetorheological finishing. J Mod Phys B 10:2857–2865
3. Seok J, Kim Y-J, Jang K-I, Min B-K, Lee S-J (2007) A study on the fabrication of curved surfaces using magnetorheological fluid finishing. J Mach Tools Manuf 47:2077–2090
4. Kim W, Lee S-H, Min B-K (2004) Surface finishing and evaluation of three-dimensional silicon microchannel using magnetorheological fluid. J Manuf Sci Eng 126:772–778
5. Cheng H-B, Yan Y, Wang Y-T (2009) Experimentation on MR fluid using a 2-axis wheel tool. J Mater Process Technol 209:5254–5261
6. Sidpara A, Jain V-K (2012) Nano level finishing of single crystal silicon blank using magnetorheological finishing. Tribol Int 47: 159–166
7. Sidpara A, Jain V-K (2012) Theoretical analysis of forces in magnetorheological fluid based finishing process. J Mech Sci 56: 50–59
8. Sidpara A, Jain V-K (2013) Analysis of forces on the free form surface in magnetorheological fluid based finishing process. J Mach Tools Manuf 69:1–10
9. A-k S, Jha S, Pandey P-M (2011) Design and development of nanofinishing process for 3D surfaces using ball end MR finishing tool. J Mach Tools Manuf 51:142–151
10. A-k S, Jha S, Pandey P-M (2012) Nano finishing of typical 3D ferromagnetic work-piece using ball end magnetorheological finishing process. J Mach Tools Manuf 63:21–31

11. A-k S, Jha S, Pandey P-M (2012) Parametric analysis of an improved ball end magnetorheological finishing process. J Eng Manuf 226(9):1550–1563
12. A-k S, Jha S, Pandey P-M (2013) Mechanism of material removal in ball end magnetorheological finishing process. Wear 302: 1180–1191
13. Jung B, Jang K, Min B, Lee S, Seok J (2009) Magnetorheological finishing process for hard materials using sintered iron-CNT compound abrasives. J Mach Tools Manuf 49(5): 407–418
14. Niranjan MS, Jha S (2013) Development of sintered magnetic abrasives and MR finishing using ball end magnetorheological finishing tool. Proceeding of international conference on precision, meso, micro and nano engineering. NIT Calicutta 536–540
15. Niranjan MS, Jha S (2013) Characterization of sintered magnetic abrasives and MR finishing using ball end magnetorheological finishing tool. Proceeding of national conference on manufacturing: vision for future, IIT Guwahati 44–52
16. Niranjan MS, Jha S (2014) Flow behaviour of bidisperse MR polishing fluid and ball end MR finishing. Proceeding of international conference on material processing and characterization, GRIET Hyderabad
17. Upadhyaya A, Upadhyaya G-S (2013) Powder metallurgy science, technology and materials. University (India) Press Pvt Ltd, India

Author's Accepted Manuscript



Fingerprint denoising using ridge orientation based clustered dictionaries

Kuldeep Singh, Rajiv Kapoor, Raunaq Nayar

www.elsevier.com/locate/neucom

PII: S0925-2312(15)00522-6
DOI: http://dx.doi.org/10.1016/j.neucom.2015.04.053
Reference: NEUCOM15418

To appear in: *Neurocomputing*

Received date: 31 December 2014
Revised date: 9 March 2015
Accepted date: 14 April 2015

Cite this article as: Kuldeep Singh, Rajiv Kapoor, Raunaq Nayar, Fingerprint denoising using ridge orientation based clustered dictionaries, *Neurocomputing*, http://dx.doi.org/10.1016/j.neucom.2015.04.053

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

Fingerprint denoising using ridge orientation based clustered dictionaries

Kuldeep Singh[a*I], Rajiv Kapoor[b], Raunaq Nayar[c]

[a]Central Research Lab, Bharat Electronics Ltd., Ghaziabad, 201010, India
[b]Dept. of Electronics & Communication, Delhi Technological University, Delhi, 110042, India
[c]Dept. of Electronics & Communication, Delhi Technological University, Delhi, 110042, India
[*]Corresponding author. Tel.: +91 991 010 1592.
kuldeep.er@gmail.com

12/31/2014; 3/9/2015; 4/14/2015

Abstract

In this paper, we introduce a novel approach of fingerprint denoising using ridge orientation based clustered sub dictionaries. The idea behind this approach is to group the patches of similar geometric structures or dominant orientation and construct separate sub-dictionaries for each group. The orientation of ridge or a valley has been exploited in finger print matching algorithms in the past. In the proposed method, the same concept of ridge orientation is utilized to group the patches and to subdivide a large dictionary into array of sub dictionaries. The new approach undergoes three steps i.e. *Ridge orientation based clustering*, *Dictionary learning* and *sparse coefficient calculation*. While reconstructing the image in final step, the minimum residual error criterion is used for choosing sub dictionary for a particular patch. The algorithm performance is experimentally compared with other existing methods not only in terms of PSNR, SSIM measures but also in terms of Euclidean distance parameter, which is used, for fingerprint matching procedures. The simulation results demonstrates that the new method achieves better results in comparison with its counterparts and will establish in improving performances of fingerprint-identification systems.

Highlights

- Use of sub dictionaries based on dominant orientation to improve the effectiveness of sparse modeling of information in a fingerprint image in the form local ridge patterns.
- Inclusion of sub dictionary for smooth patches improves the reconstruction where, no dominant ridge orientation is present.
- Adaptive selection of suitable sub dictionary while reconstruction based on minimum residue error criterion enhances the performance of algorithm.

Accepted manuscript

ACCEPTED MANUSCRIPT

Keywords

Ridge Orientation; Sparse Coding; Euclidean Distance; Fingerprint

1. Introduction

Automatic fingerprint identification systems (AFIS) are integral part of man machine interface systems for verification of individual's identity. Fingerprints has proved to be an effective biometric identifier and extensively utilized in forensic science for criminal investigations [1] and in biometric identification devices for civilian as well as commercial applications [2]. Efficiency of fingerprint recognition systems is largely dependent on the input fingerprint image quality [3]. There could be various factors behind the substandard quality of fingerprint images, such as performance of the image acquisition sensor, image corruption with noise due to variation in skin and impression condition and so on. Probability of error in the recognition process can be reduced by quality improvement through image denoising.

Image denoising is a challenging problem and has acquired considerable interest of researchers from both academia and industry. The classic image denoising problem is modeled by assuming an ideal image X being captured in the presence of an additive zero-mean white and homogeneous Gaussian noise with standard deviation σ. The captured image Y is represented by,

$$Y = X + n\left(\sigma^2\right) \quad (1)$$

There are numerous denoising algorithms [4–9] present in literature to estimate X from Y. Here we are focusing towards one specific approach i.e. sparsity based image denoising. Elad and Aharon [10] addressed image Denoising problem using K-SVD algorithm [11] of optimal overcomplete dictionary construction. In this method each patch is denoised by representing each patch as a linear combination of only few atoms in the dictionary. Mairal et al. [12] proposed multiscale global dictionaries for performance improvement of dictionary learning. Chatterjee and Milanfar [13] proposed locally learned dictionaries (K-LLD) for image denoising, which performs clustering based on similar geometric structures and employs local weight functions as features. Xiaoqiang et al. [14] proposed a Bayesian-based sparse coding algorithm, which employs spike and slab prior to provide accurate prediction and effective sparse representation. Zhang et al. [15] introduced a new approach of image denoising, which creates 2D dictionary based on the self-similarity inherent in the images. This method relies on idea that if group of similar patches are assembled in matrix form, then there exist linear correlations among both columns and rows. Yang et al. [16] proposed new dictionary learning for image noise reduction, which exploits structural similarities of image. Dabov et al. [17] proposed a state-of-art method, which yields

Accepted manuscript

excellent PSNR performance. BM3D algorithm is a unified framework of dictionary learning and structural clustering.

The problem of fingerprint denoising is peculiar in nature and need special methods for denoising. In this paper, we propose a novel framework for denoising fingerprint images. Instead of having a single dictionary, the proposed method clusters the training patches into groups based on dominant orientation of patch. Authors believe that use of sub dictionaries based on dominant orientation best describe the underlying image data and improves the effectiveness of sparse modeling of information in a fingerprint image in form of local ridge patterns. Atoms from sub dictionary with minimum residual error are used during reconstruction of patches.

The rest of this paper is organized as follows. In Section 2, the sparse reconstruction model is depicted. Section 3 describes the proposed fingerprint denoising algorithm. In Section 4, experimental results are presented to evaluate the performance and demonstrate the effectiveness of our proposed method. Finally, conclusions are presented in Section 5.

2. Sparse Reconstruction Model

The concept of sparse based image denoising is motivated from the idea that a high quality image *X* can be approximately coded as a linear combination of few columns (atoms) of an over complete dictionary. Overcomplete dictionaries contains more atoms than the dimension of the signal. A signal $X = \{x_1, x_2 \ldots..., x_n\}$ is said to be sparse when very few of entries x_i possess non-zero values. Sparse representation computes the summation of the constituent atoms weighted with their sparse coefficients vector. However, computing such sparse codes within overcomplete dictionaries is non-trivial, in particular because the decomposition of an image in terms of atoms of an overcomplete dictionary is not unique. Processing the whole image as a large vector is numerically cumbersome, Elad & Aharon (2006) [10], proposed breaking down the image into smaller patches and learning a dictionary of patch-sized atoms. Handling high dimensionality of data in image denoising is major challenge, the patch based method models the patches with relatively lower dimension, seems the effective solution to tackle this problem.

Given an image observation patch of size $\sqrt{p} \times \sqrt{p}$ ordered lexicographically as column vector $x \in R^p$ and a sparsifying dictionary $\varphi = \{\varphi_1, \varphi_2 \ldots..., \varphi_c\} \in R^{n \times M}$ (the columns $\varphi_i \in R^M$ represent the atoms), sparse representation refer to finding an coefficient vector $\alpha \in R^p$ in the domain spanned by the dictionary φ while synthesizing following equation

Accepted manuscript

$$\hat{\alpha} \in \text{argmin}_{\alpha} \| \alpha \|_0 \; subject\ to\ x = \varphi\alpha \quad (2)$$

The solution is indeed very sparse where $\| \hat{\alpha} \|_0 << n$. While synthesizing approximately, the equality constraint in Eq. (2) can be replaced by a l_2 norm inequality constraint.

$$\hat{\alpha} \in \text{argmin}_{\alpha} \| \alpha \|_0 \; subject\ to \| x - \varphi\alpha \|_2 \leq \varepsilon_1 \quad (3)$$

where ε controls the misfitting between observed and recovered signal $\hat{x} = \varphi\hat{\alpha}$. Now considering a patch y, noisy version of x, contaminated with additive white Gaussian noise. The denoising of noisy patch y can be formulated as following optimization problem.

$$\hat{\alpha} \in \text{argmin}_{\alpha} \| \alpha \|_0 \; subject\ to \| y - \varphi\alpha \|_2^2 \leq \varepsilon_2 \quad (4)$$

The denoised patch can be recovered as $\hat{x} = \varphi\hat{\alpha}$. The optimization problem in Eq. (4) can also be reformulated in its unconstrained penalized form.

$$\hat{\alpha} = \text{argmin}_{\alpha} \| y - \varphi\alpha \|_2^2 + \beta \| \alpha \|_0 \quad (5)$$

The solution of Eq. (5) is NP hard problem; however, l_1-norm can be used to make it computationally tractable (convex) as an alternate to l_0-norm (non-convex). Matching pursuit class of algorithms that include orthogonal matching pursuit (OMP), basis pursuit (BP) and gradient pursuit (GP) provide efficient and computationally tractable solutions to the optimization problem.

3. Proposed Method

3.1. Sparse based denoising framework

The sparse based denoising framework comprises of restoring true image X from its noisy version Y contaminated with additive white Gaussian noise. The task involves suitable dictionary learning and coefficient calculation for efficient sparse representation of denoised image X. Elad and Aharon [10] proposed an image-denoising framework, which performs alternate minimization for learning patch-based dictionary where the size of each atom is equal to patch size. The energy minimization problem for image denoising boils down to

$$\{\hat{\alpha}, \hat{X}\} = \text{argmin}_{\alpha_{i,j}, X} \gamma \| X - Y \|_2^2 + \sum_{i,j} (\| \varphi\alpha_{i,j} - P_{i,j}(X) \|_2^2 + \beta \| \alpha_{i,j} \|_0) \quad (6)$$

here $P_{i,j}(X)$ is patch extraction operator. The above minimization process is combination of three penalty terms; the first term imposes demand of proximity between the measured image,

Accepted manuscript

and its denoised counterpart. The second and third term demands that each reconstructed patch can be sparsely represented as linear combination of very few coefficients and dictionary with minimum error. The image denoising algorithm via energy minimization in Eq. (6) involves an iterative block-coordinate relaxation method, that updates sparse coefficients, dictionary and image alternatively once at a time and keep other unknowns fixed.

3.2. Ridge orientation dictionary based image denoising

The denoising algorithm comprises of three main steps: Ridge orientation based clustering, Dictionary learning and sparse coefficient calculation.

3.2.1. Ridge orientation based clustering

A fingerprint is the pattern of ridges and valleys on the surface of the finger. Considering the fact that images especially fingerprint images often contains many repetitive geometric structures (or self-similarity) such as ridges and valleys, we exploit this self-similarity among patches for better sparse representation, which in turn increases the denoising performance. The idea behind this approach is to group the patches containing similar geometric structures or dominant orientation and construct separate sub-dictionaries for each group. The orientation of ridge or a valley has been exploited in finger print matching algorithms [18,19] in past. In the proposed method, we have utilized the same concept of ridge orientation in fingerprint to group the patches and to subdivide a large dictionary into array of sub dictionaries. Ridge and valleys are main texture information embedded in fingerprint images and sparse modeling based on these features can significantly enhance the performance of sparse representation based denoising approach. Firstly, the patches are categorized into smooth and rough patch by applying a variance threshold τ on each image patch. The rough patches are further classified into dominant and non-dominant orientation patches by applying threshold on the dominant measure calculated by singular value decomposition (SVD) of gradient vector of each patch. The estimation of local orientation is based on singular value decomposition (SVD) of the gradient vectors of the patch [20]. The method for this classification is described in following:

Calculate the local estimate of gradient $\nabla y(u,v)$ at each i^{th} pixel at location (u,v) in the image patch.

$$\nabla y(u,v)_i = \left\{ \frac{\partial I(u,v)}{\partial u}, \frac{\partial I(u,v)}{\partial v} \right\} \quad (7)$$

The gradient map $G \in R^{N\times 2}$ of the $\sqrt{p}\times\sqrt{p}$ patch is denoted as

$$G = \{\nabla y(u,v)_1, \nabla y(u,v)_2, ..., \nabla y(u,v)_p\}^T \quad (8)$$

Accepted manuscript

ACCEPTED MANUSCRIPT

Next step is to compute SVD of the gradient map

$$G = USV^T \quad (9)$$

here $U \in R^{N\times N}$ represents each vector's contribution to the corresponding singular vector; $S \in R^{N\times 2}$ represent the energy in the dominant directions; and $V \in R^{2\times 2}$ represent orientation, in which the columns v1 and v2 represents the dominant and subdominant orientation of the gradient field respectively. The categorization of rough patches into dominant or non-dominant orientation is decided based on a dominant measure *R* calculated as

$$R = \frac{S_1 - S_2}{S_1 + S_2} \quad (10)$$

here S_1 and S_2 are singular values representing the energy value in the dominant direction. The patches having dominant measure lesser than the significance level threshold R^* i.e. the singular values are remarkably different, it is very likely that the corresponding patch is only pure white noise and contains no dominant orientation [20]. The rough patches are categorized in dominant or non-dominant orientation patches for which the value of *R* is higher or lesser than significance level threshold R^* respectively.

The ridge orientation angle of all the dominant oriented patches is estimated by horizontal and vertical gradients of the patch. For finding ridge orientation angle of each patch, a gradient operator, is applied to obtain the horizontal gradient value G_r and vertical gradient value G_s and the block horizontal and vertical gradients i.e. G_{rr} and G_{rs}, are obtained by adding up all the pixel gradients of the corresponding direction.

$$G_{rs} = \sum_{u=i-\left(\frac{w}{2}\right)}^{i+\left(\frac{w}{2}\right)} \sum_{v=j-\left(\frac{w}{2}\right)}^{j+\left(\frac{w}{2}\right)} 2G_r(u,v)G_s(u,v) \quad (11)$$

$$G_{rr} = \sum_{u=i-\left(\frac{w}{2}\right)}^{i+\left(\frac{w}{2}\right)} \sum_{v=j-\left(\frac{w}{2}\right)}^{j+\left(\frac{w}{2}\right)} (G_r^2(u,v) - G_s^2(u,v)) \quad (12)$$

The most probable ridge orientation of the patch can now be obtained as

$$O(u,v) = \frac{1}{2}\tan^{-1}\left(\frac{G_{rs}}{G_{rr}}\right) \quad (13)$$

Accepted manuscript

Once the ridge orientation of each dominant oriented patch is estimated further clustering is applied on the angles to group patches of similar structure in one cluster. For our algorithm, we run unsupervised classification technique, K-means clustering algorithm, which performs clustering of patches into some prefixed number of clusters. At the end of the clustering process, we have *N* clusters (classes Ω_n), one for smooth, and one for non-dominant orientation patches and $N-2$ clusters for different ridge orientation angle patches.

3.2.2. Dictionary learning

The image divided in overlapped patches of size $\sqrt{p}\times\sqrt{p}$ and each patch is ordered lexicographically as column vector $y_k \in R^p$; here k is index of patch. Next all the patches belonging to same sub group are stacked in matrix denoted as $y_{\Omega_n} \in R^{p\times C_n}$ i.e. $y_{\Omega_n} = \{y_{\Omega_{n1}}, y_{\Omega_{n2}}, \ldots\ldots, y_{\Omega_{nC}}\}$, which consists of each patch represented by column of matrix. C_n is the number of patches in sub groups and it may vary for different group. Once groups of patches of similar orientation are formed, next step is to train sub dictionary for individual group. The core of the sparse representation modeling lies in the choice of dictionary φ.

The goal of dictionary learning is to optimize the dictionary and sparse coefficients jointly utilizing a set of given training image patches. The sub dictionary training for each cluster can be formulated as

$$\{\hat{\alpha}_{\Omega_n}, \varphi^{(\Omega_n)}\} = \text{argmin}_{\alpha,\varphi} \sum_{k\in\Omega_n} (\| \varphi^{(\Omega_n)} \alpha_{\Omega_n} - y_k \|_2^2 + \beta \| \alpha_{\Omega_n} \|_0) \quad (14)$$

here $n = 1, 2, \ldots..N$ is index for sub dictionaries learned for each sub group of patches classified based on dominant orientation. K-SVD [11] is natural choice for dictionary learning due to its simplicity and ease of implementation. In addition to it, the K-SVD dictionary learning process has in it a noise rejection capability [10]. The MATLAB code of the KSVD algorithm for dictionaries learning is taken from http://www.cs.technion.ac.il/~elad/software. The dictionary learning step yields *N* sub dictionary $\{\varphi^{(\Omega_1)}, \varphi^{(\Omega_2)}, \ldots.., \varphi^{(\Omega_N)}\}$ and number of atoms are adaptive to the total number of patches in each group. For fingerprint images, the smooth and non-dominant orientation groups have less number of patches; hence, lesser number of atoms can be allotted to corresponding sub dictionary. Fig. 1 shows typical sub dictionaries learned based on ridge orientation based clustered patches.

Accepted manuscript

3.2.3. Sparse coefficient calculation

Once the sub dictionaries are learned for each cluster, the next step is to calculate the sparse coefficients to obtain the denoised image. Estimation of the sparse coefficients $\hat{\alpha}$, comes down to solve Eq. (5).Orthogonal Matching Pursuit (OMP) algorithm is best choice due to its speed and easy parameter tuning facility for solution of given optimization equation. Given the set of sub dictionary $\{\varphi^{(\Omega_1)}, \varphi^{(\Omega_2)}, \ldots\ldots, \varphi^{(\Omega_N)}\}$ we form sparse coefficient matrix for each patch corresponding to every sub dictionary. The Eq. (5) can be re-formulated as

$$\hat{\alpha}^{(\Omega_n)} = \text{argmin}_{\alpha} \| y_k - \varphi^{(\Omega_n)} \alpha \|_2^2 + \beta \| \alpha \|_1 \quad (15)$$

here $n = 1, 2, \ldots\ldots N$ is index for sub dictionaries.

To recover the denoised patch for a given noisy patch, the sparse coefficients $\{\hat{\alpha}^{(\Omega_n)}\}$ are calculated using each of the sub dictionaries and the residual error $\epsilon^{(\Omega_n)}$ is calculated by following equation for each dictionary

$$\left\{\epsilon^{(\Omega_n)}\right\} = \| y_k - \varphi^{(\Omega_n)} . \hat{\alpha}^{(\Omega_n)} \|_2 \quad (16)$$

The sub dictionary having least residual error is chosen for that particular patch for reconstruction of corresponding denoised patch. The final denoised image is obtained by aggregation of all estimated denoised patches $\{y_k\}$.

4. Experimental Results

To demonstrate the performance of proposed method we have devised a validation methodology consisting of three experiments. In the first experiment, the proposed method is compared with NLM [21], K-SVD [10] and KLLD [13] methods in terms of quality measures i.e. PSNR and SSIM. Further, the proposed method is investigated in terms of visual quality. Authors believe that visual inspection and quality measures are not sufficient to evaluate the performance of fingerprint denoising. Filter bank based fingerprint matching [18] techniques have proved to be very effective in fingerprint matching. We have also used the same concept for performance evaluation and compared the Euclidean distance parameter of proposed method with other methods. Zero mean additive white Gaussian noise with different standard deviation was added in images to produce noisy version. The effectiveness of the proposed method is evaluated on public databases FVC2000, FVC2002 and FVC2004. The proposed algorithm has been implemented in MATLAB R9, running on a Windows PC with Intel Core i5 at

Accepted manuscript

ACCEPTED MANUSCRIPT

2.60 GHz and 4 GB of RAM. We have set empirical values of variance threshold $\tau = 50$ and significance level threshold $R^* = 0.175$ for the classification of image patches into smooth, dominant or non-dominant orientation. These values can be manually fine-tuned as per requirement of the data set, taking higher or lower values of these parameters can result into classification of smooth/non-dominant orientation patch into dominant orientation patch and vice versa. Total seven sub dictionaries including one for smooth and one for non-dominant orientation are learned from the group of patches clustered on the basis of ridge orientation. Based on the distribution of patches in each group number of clusters is chosen. Higher number of clusters is not providing any significant improvement in results and sometimes results into very few patches in a sub group.

4.1. Comparison of denoising performance

For assessment of performance of the algorithm, we have chosen two quality measures named as Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise and expressed in terms of the logarithmic decibel scale. The PSNR is calculated as

$$PSNR = 10\times \log\frac{255^2}{MSE}; MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[A(m,n)-B(m,n)]^2 \quad (17)$$

A is original image and B is output of proposed method.

SSIM measures the structural similarity between the input and output images. SSIM depicts the distortion of the visual sensing. The SSIM is defined as

$$SSIM(a,b) = \frac{(2\mu_a\mu_b + c_1)(2\sigma_{ab} + c_2)}{(\mu_a^2 + \mu_b^2 + c_1)(\sigma_a^2 + \sigma_b^2 + c_2)} \quad (18)$$

where μ_a and μ_b are the mean values of images A and B, σ_a and σ_b are the standard deviation, σ_{ab} is the covariance, $c_1 = (K_1L)^2$ and $c_2 = (K_2L)^2$,with L being the maximum pixel value, K_1=0.01 and K_2=0.03) [22].

Observation noise with varying variance is added and four typical values of sigma 10, 20, 30 and 40 are taken for experimentation. The PSNR values depicted in Table 1 clearly defines the superiority of proposed method over other methods. Dataset of 40 test images selected from three databases is constructed and mean PSNR is calculated for all the methods and results are presented in Table 1 for

Accepted manuscript

ACCEPTED MANUSCRIPT

analysis purpose. From the analysis of results, we hereby conclude that the proposed algorithm has achieved highly competitive PSNR performance for all the target images.

Similar to Table 1 mean SSIM is calculated on same dataset and Table 2 depicts the SSIM performance results of proposed method in comparison with other three methods. From the comparative analysis, it is concluded that the proposed algorithm provides excellent SSIM performance for higher noise levels.

4.2. Performance assessment based on fingerprint matching

Euclidean distance parameter [18] is used here for evaluating the performance in terms of finger print matching. For this, we have used MATLAB code available at http://www.advancedsourcecode.com/fingerprint.asp. Results in Table 3 clearly indicate that proposed method is superior in terms of fingerprint matching. The Euclidean distance for proposed method is significantly less than other methods for most of the cases.

4.3. Visual performance comparison

Figs. 2–5 show the visual performance comparison of proposed method with KLLD, KSVD and NLM method. It is important to emphasize that the visual impressions of resultant image of proposed method are very close to the original image. The visual similarity is the key feature for fingerprint matching; therefore, it is an important aspect to be taken under consideration while performing image denoising of fingerprints. From the results, it is evident that KSVD and NLM methods have introduced distortion of blur in the fingerprints in Fig. 3 and Fig. 4, however the KLLD results are also close to the original image but KLLD is not able to filter out the noise completely. The visual results accord with the numerical results in terms of Euclidean distance.

5. Conclusion

We have proposed a new algorithm specific to fingerprint denoising. Instead of having a single dictionary, the proposed method clusters the training patches into groups based on ridge orientation of patch. Since every sub-group consists of geometrically similar patches, thus, the sub-dictionary trained from it can work more effectively and this will be help for improving denoising performance. Use of sub dictionaries based on dominant orientation significantly improves the effectiveness of sparse modeling of information in a fingerprint image in the form local ridge patterns. Experimental results

Accepted manuscript

ACCEPTED MANUSCRIPT

demonstrate that the proposed approach can be very helpful in denoising of fingerprints and can provide aid to automatic fingerprint identification systems.

References

[1] M.D. Marsico, M. Nappi, G. Tortora, NABS: Novel approaches for biometric systems, IEEE Trans. Syst.,Man, Cybern. Appl. Rev. 41(4) (2011) 481–493.

[2] A.K. Jain, A. Ross, S. Pankanti, Biometrics: A tool for information security, IEEE Trans. Inf. Forensics Security 1(2) (2006) 125–143.

[3] Weisheng Dong, L.e.i. Zhang, Guangming Shi and Xiaolin Wu, Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization, IEEE Trans. on Image Process. 20 (7) (2011) 1838-1857.

[4] S. Grace Chang, Bin Y.U., and Martin Vetterli, Adaptive Wavelet Thresholding for Image Denoising and Compression, IEEE Trans. on Image Process. 9(9) (2000) 1532-1546.

[5] M.K. Michak, Igor Kozintsev, Kannan Ramchandran, Pierre Moulin, Low-Complexity Image Denoising Based on Statistical Modeling of Wavelet Coefficients, IEEE Signal Processing Letters 6 (12) (1999) 300-302.

[6] L. Sundur, I.W. Selesnick, Bivariate Shrinkage Functions for Wavelet-Based Denoising Exploiting Interscale Dependency, IEEE Trans. on Signal Process. 50(11) (2002) 2744-2756.

[7] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image Denoising Using Scale Mixtures of Gaussian in the Wavelet Domain, IEEE Trans. on Image Process. 12(11) (2003) 1338-1351.

[8] J.i.n. Wang, Yanwen Guo, Yiting Ying, Yanli Liu, Qunsheng Peng, Fast non-local Algorithm for Image Denoising, in: Proceedings of IEEE International conference on Image Processing, 2006, 1429-1432.

[9] Mark Miller, Nick Kingsbury, Image Denoising Using Derotated Complex Wavelet Coefficients, IEEE Trans. Image Process. 17(9) (2008) 1500-1511.

[10] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process. 15(12)(2006) 3736–3745.

[11] M. Aharon, M. Elad, and A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Trans. Signal Process., 54, no. 11, 4311–4322, Nov. 2006.

[12] J. Mairal, G. Sapiro, M. Elad, Learning multiscale sparse representations for image and video restoration, SIAM Multiscale Model. Simul. 7(1)(2008) 214–241.

Accepted manuscript

ACCEPTED MANUSCRIPT

[13] P. Chatterjee, P. Milanfar, Clustering-Based Denoising With Locally Learned Dictionaries, IEEE Trans. Image Process. 18 (7) (2009) 1438–1451.

[14] Xiaoqiang Lu, Yuan Yuan, Pingkun Yan, Sparse coding for image denoising using spike and slab prior, Neurocomputing 106 (2013) 12–20.

[15] Xuande Zhang, Xiangchu Feng, Weiwei Wanga, Guojun Liu, Image denoising via 2D dictionary learning and adaptive hard thresholding, Pattern Recognition Letters 34 (2013) 2110–2117.

[16] Shuyuan Yang, Linfang Zhao, M.i.n. Wang, Yueyuan Zhang and Licheng Jiao, Dictionary learning and similarity regularization based image noise reduction, J. Vis. Commun. Image R. 24 (2013) 181–186.

[17] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 16 (2007) 2080–2095.

[18] A.K. Jain, S. Prabhakar,L. Hong,S. Pankanti, Filter bank-based fingerprint matching, IEEE Trans. Image Process 9 (5) (2000) 125–143

[19] A.K. Jain, S. Prabhakar, and S. Pankanti, A Filterbank-based Representation for Classification and Matching of Fingerprints, in: Proceedings of International Joint Conference on Neural Networks, 1999, 3284-3285.

[20] X.G. Feng and P. Milanfar, Multiscale principal components analysis for image local orientation estimation, in: Proceedings of 36th Asilomar Conf. Signals, Syst. Comput., Nov. 2002, 478–482.

[21] A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in: Proceedings of IEEE Conf. Computer Vision and Pattern Recognition, Oct. 2005, 60–65.

[22] Z..Wang, A.C..Bovik, H.R..Sheikh, E.P..Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

Accepted manuscript

Kuldeep Singh received his B.E. degree in Electronics & Comm. from Govt. Engg. College, Ajmer India in 2004 and M. Tech. degree in Signal Processing from Delhi University in 2006.He is currently pursuing PhD from Delhi Technological University, India. He currently works as scientist at Bharat Electronics Ltd. India. His research interest includes image enhancement, sparse representation based image processing and pattern recognition. He is also reviewer of various Elsevier and IET journals.

Rajiv Kapoor received his B.E., M.E. degree from Delhi University, India and PhD from Punjab University, India in the field of Electronics & Communication.He is currently Professor at Deptt. of Electronics & Comm. Delhi Technological University, Delhi, India. He is also Editor of ST Micro

ACCEPTED MANUSCRIPT

Electronics International Journal in the field of Electronics Design. He has authored more than 45 research papers. He is also reviewer of various IEEE, IET, Elsevier and Springer Journals. His research interest includes Image Processing, Object Tracking, Pattern Recognition and Character Recognition.

Raunaq Nayar his engineering under graduate student from Delhi Technological University in the field of Electronics & Communication. He is a research enthusiast especially in the field of Image processing and Image quality enhancement.

Fig. 1 Ridge orientation based clustered sub dictionaries (a) to (e) Dominant orientation (f). Smooth (g). Non-dominant orientation.

Fig. 2 Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method.

Fig. 3 Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method.

Fig. 4 Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method.

Fig. 5 Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method.

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 1 The PSNR results (in dB) of all algorithms for various noise levels.

Image	Sigma	Proposed Method	KSVD	NLM	KLLD
Fingerprint 1	10	**31.6769**	30.9895	31.3191	30.8434
	20	**27.7703**	27.5687	27.1032	27.7517
	30	25.5703	**25.8326**	25.6037	24.5308
	40	**24.6637**	24.6077	24.2337	21.7344
Fingerprint 2	10	**30.4791**	30.2634	30.3297	29.9843
	20	**26.8613**	26.4305	26.5142	26.2517
	30	**24.6845**	24.6765	24.5067	24.4398
	40	23.5657	**23.7767**	23.4236	22.9374
Fingerprint 3	10	32.3242	32.0569	**32.5172**	32.1924
	20	**28.8809**	28.4793	28.6791	28.7545
	30	**26.7896**	26.6678	26.7243	26.6304
	40	25.2794	25.3067	**25.3372**	24.6609
Fingerprint 4	10	**31.7684**	31.4329	31.5460	30.9760
	20	**27.8767**	27.6472	27.7545	27.1518
	30	**25.7345**	25.6985	25.4057	25.2378
	40	24.5574	**24.5796**	24.3289	24.1384
Mean (40 Images)	10	**32.5621**	32.0856	32.3280	31.9990
	20	**26.8473**	26.4314	26.1127	26.0774
	30	**25.6947**	25.1188	25.5601	25.2097
	40	**23.5165**	23.3676	23.3308	22.3677

Table 2 The SSIM results of all algorithms for various noise levels.

Image	Sigma	Proposed Method	KSVD	NLM	KLLD
Fingerprint 1	10	**0.9210**	0.8807	0.8761	0.9035
	20	0.8337	0.8026	0.8126	**0.8594**
	30	**0.7856**	0.7690	0.7791	0.7674
	40	**0.7680**	0.7421	0.7537	0.6634
Fingerprint 2	10	**0.9323**	0.9007	0.9167	0.9129
	20	0.8554	**0.8664**	0.8517	0.8429
	30	**0.8119**	0.8097	0.8023	0.8009
	40	**0.7878**	0.7756	0.7796	0.7239
Fingerprint 3	10	**0.9192**	0.9118	0.9087	0.9034
	20	**0.8426**	0.8336	0.8419	0.8287
	30	**0.7998**	0.7726	0.7829	0.7817
	40	**0.7678**	0.7624	0.7569	0.7556
Fingerprint 4	10	0.8947	**0.8959**	0.8897	0.8761
	20	**0.8267**	0.8213	0.8179	0.8127
	30	**0.7786**	0.7623	0.7645	0.7527
	40	**0.7492**	0.7337	0.7423	0.7242
Mean (40 Images)	10	**0.9168**	0.8772	0.8678	0.8589
	20	**0.8396**	0.8209	0.8110	0.8059
	30	**0.7939**	0.7784	0.7822	0.7756
	40	**0.7682**	0.7434	0.75812	0.7167

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 3 The Euclidean distance results of all algorithms for various noise levels.

Image	Sigma	Proposed Method	KSVD	NLM	KLLD
Fingerprint 1	10	**20.3062**	22.3375	115.6792	26.5409
	20	**60.9436**	88.9767	98.7143	74.2789
	30	**91.1935**	129.8676	133.5734	98.5879
	40	**100.0349**	136.6567	126.5106	111.6721
Fingerprint 2	10	**32.3062**	36.8976	69.7865	33.2789
	20	87.8761	87.6678	98.8978	**86.2639**
	30	**102.3498**	103.5610	112.5678	121.7754
	40	**121.6123**	124.1245	123.2730	125.7699
Fingerprint 3	10	**26.7645**	29.3321	26.8961	27.8510
	20	**56.5643**	57.2331	59.0089	58.9002
	30	**80.0126**	82.2094	80.9429	83.2184
	40	**98.0197**	99.2970	100.5697	100.6573
Fingerprint 4	10	40.5692	**40.1297**	48.0371	40.3201
	20	**85.9709**	89.3264	86.2340	88.0362
	30	105.6931	109.3287	109.2367	**103.1017**
	40	**132.0289**	133.3079	138.4563	137.2843

Accepted manuscript

ACCEPTED MANUSCRIPT

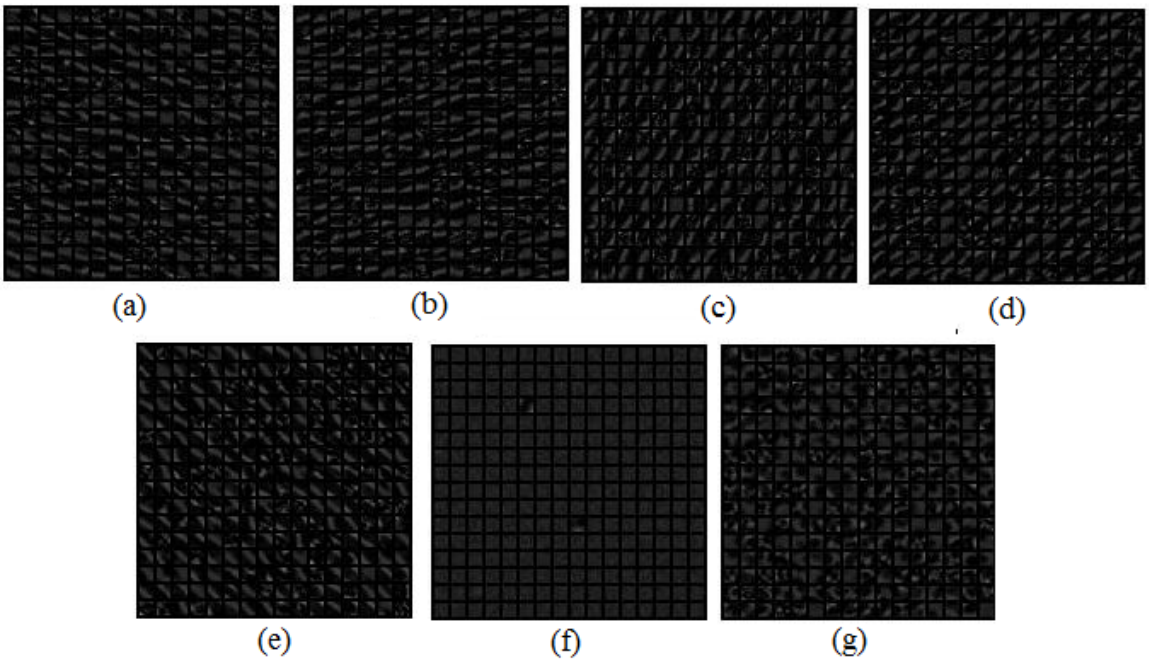


Fig.1. Ridge orientation based clustered sub dictionaries (a) to (e) Dominant orientation (f). Smooth (g). Non-dominant orientation

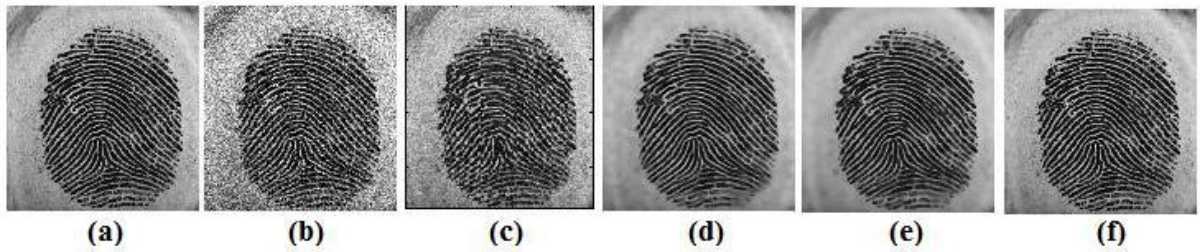


Fig.2. Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method

ACCEPTED MANUSCRIPT

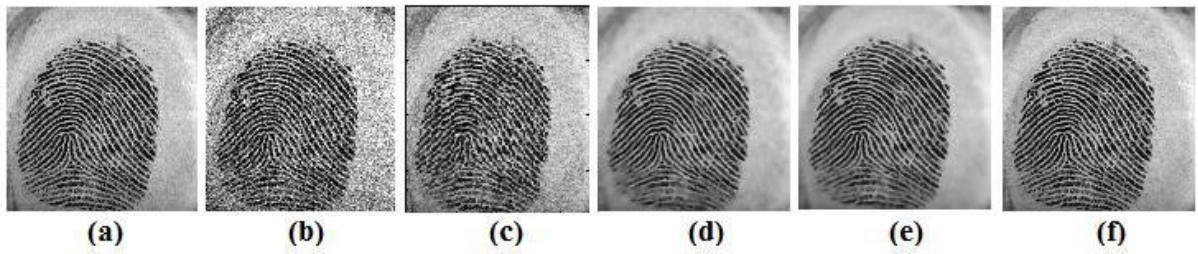


Fig.3. Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method

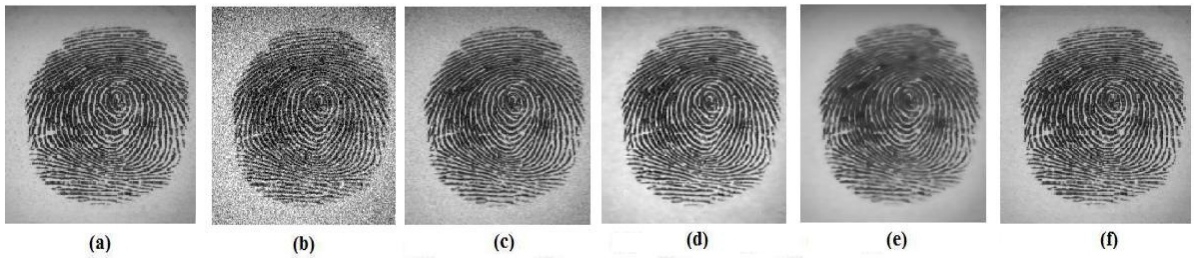


Fig.4. Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method

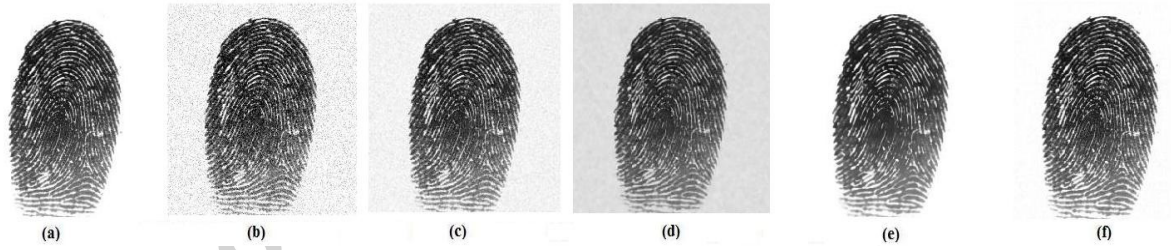


Fig.5. Visual performance comparison of different methods (a). Original (b). Noisy (c). K-LLD (d). K-SVD (e). NLM (f). Our Method

*Biography of the author(s) ACCEPTED MANUSCRIPT

BIOGRAPHIES

Kuldeep Singh received his B.E. degree in Electronics & Comm. from Govt. Engg. College, Ajmer India in 2004 and M. Tech. degree in Signal Processing from Delhi University in 2006.He is currently pursuing PhD from Delhi Technological University, India. He currently works as scientist at Bharat Electronics Ltd. India. His research interest includes image enhancement, sparse representation based image processing and pattern recognition. He is also reviewer of various Elsevier and IET journals.

Rajiv Kapoor received his B.E., M.E. degree from Delhi University, India and PhD from Punjab University, India in the field of Electronics & Communication .He is currently Professor at Deptt. of Electronics & Comm. Delhi Technological University, Delhi, India. He is also Editor of ST Micro Electronics International Journal in the field of Electronics Design. He has authored more than 45 research papers. He is also reviewer of various IEEE, IET, Elsevier and Springer Journals. His research interest includes Image Processing, Object Tracking, Pattern Recognition and Character Recognition.

Raunaq Nayar his engineering under graduate student from Delhi Technological University in the field of Electronics & Communication. He is a research enthusiast especially in the field of Image processing and Image quality enhancement.

Accepted manuscript

*Photo of the author(s)

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



*Photo of the author(s)

*Photo of the author(s)

ACCEPTED MANUSCRIPT



International Journal of Advanced Information Science and Technology (IJAIST) ISSN:2319:2682
Vol.36, No.36, April 2015

Fuzzy Modelling For Selection Of Most Economical Green Building Out Of N-Alternatives

Sunita Bansal
Associate Professor
Department of Civil Engineering
Manav Rachna International University
Faridabad, India
E-mail : sunitabansal.fet@mriu.edu.in

Prof. (Dr.) Srijit Biswas, FIE
Department of Civil Engineering
Manav Rachna International University
Faridabad, India
E-mail: srijitbiswas@yahoo.com

Prof. (Dr.) S. K. Singh
Department of Environmental Engineering
Delhi Technological University
Delhi, India
E-mail: sksinghdce@gmail.com

Corresponding Author *
Sunita Bansal
Associate Professor
Department of Civil Engineering
Manav Rachna International University
Faridabad, India
E-mail : sunitabansal.fet@mriu.edu.in

***Abstract*—Building Siting decisions are long term decisions which have to be taken considering environmental factors, owners' desires and economics of different cities. A Green Building(GB) is a multifaceted concept and is realized only, when all elements of the built environment are looked at holistically, which saves water and energy, generates less waste, provides healthier spaces with good sanitary conditions etc. out of many. But to define an economical green building in its essence out of n-alternatives, it is essential to assess the total environmental impact values (TEIV) of all buildings (inside and outside) and then optimise with respect to their constructional cost & cost of land. But it is not an easy job because most of the data are not always crisp or numeric but rather linguistic and hedges like 'high reflective roof coating', 'very bad orientation', 'poor sanitation', 'very good environmental quality', 'bad quality materials', 'good drainage system', 'heavy rainfall', 'high energy consumption', etc. All these data are fuzzy in nature thus uncertainty is the integral part of evaluation. All experts' perception towards giving their decision depends wholly on their neural network functions which fluctuate according to the nature of function of dendrite and axon. Thus every decision-maker hesitates more or less, on every evaluation activity which can be solved by using fuzzy logic. This paper presents a fuzzy model to tackle the uncertainty more precisely in selection of the most economical Green Building (GB) out of n-alternatives.**

***Keywords*: fuzzy decision, green building, Intuitionistic fuzzy sets, mean fuzzy set, etc.**

I. INTRODUCTION

For construction of the most economical green building, main concept should always be borne in mind that the building has to be energy efficient and the structure should help us to save money[6]. That means to maximize the functionalities and efficiency of an economical green building the use of local healthy materials with low cost of land are also same important like as augmentation of all positive environmental impacts[1,5]. In practice, it is often found that a building of a particular city may be energy and environmentally efficient but may not be declared as economical and vice versa even though both may have same plan, design & orientation. So it is very difficult for an expert to take decision for selection of most economical green building precisely out of n-alternatives. Because every decision maker hesitates more or less on every evaluation activity because some part of the evaluation contributes to truthness and some part of the evaluation contributes to falseness. Thus uncertainty plays a major role in the expert's perception which can be minimised only using fuzzy logic of Prof. Latfi Zadeh [7] and it's higher order fuzzy logic of Prof. K.T.Atanassov. According to fuzzy logic, a fuzzy set A is defined as a set of ordered pairs like, $A = \{(x_1, \mu_A(x_1)), (x_2, \mu_A(x_2)), \ldots\ldots , (x_n , \mu_A(x_n))\}$, where $\mu_A(x_n)$ is the degree of truth membership value of element x_n and $[1-\mu_A(x)]$ is the false membership value $[v_A (x)]$ [3,7]. But in many cases of judgement, the fuzzy logic is not sufficient to tackle the uncertainty and to overcome this problem, Prof. K. T. Atanassov introduced a new concept of Intuitionistic Fuzzy Sets (IFS) by bifurcating the Zadeh's false membership value into two separate parts i.e purely false $[v_A(x)]$ and purely indeterministic or hesitation $[\pi_A(x)]$ so that $[\mu_A(x) + v_A (x) + \pi_A (x)] = 1$ [2]. In this paper we use composed fuzzy model of IFS and Fuzzy Decision to minimise the uncertainty involved in the selection of most economical Green Building out of alternatives.

II. METHODOLOGY

Suppose n be the number of green buildings of same design and plan in the city of CT_1, CT_2, CT_3, ... CT_n, whose

total environmental impact value (TEIV) are assessed as E_1, E_2, E_3,....., E_n respectively. Suppose cost of land and constructional costs of building are not same for all buildings. To select the most economical green building out of n-alternatives, the TEIV and cost of building & land has a great role to influence the uncertainty in perception of decision makers or experts. To eliminate this uncertainty, here we use 'Intuitionistic Fuzzy Sets Theory (IFS)' of Prof. K.T. Atanassov [2] and Fuzzy Decision (FD) with a very logical treatment for overall assessment of the project. For assessment of all fuzzy data associated with evaluation process, the following definitions are essential.

A. Definition 2.1 Attributes of the Assessment

Attributes are fuzzy data for which membership value of either degree of truthness or degree of falseness lying within [1,0] is collected from every expert. As for example, if we consider a project "Assessment of 'TIEV' of a Green Building (GB)", then some relevant attributes could be 'high reflective roof coating', 'bad orientation', 'poor sanitation', 'good indoor environmental quality', 'low constructional cost', 'good drainage system', 'heavy rainfall', 'high energy consumption', 'poor insulation', etc. out of many.

B. Definition 2.2 Universe of the Assessment

Collection of all attributes of the assessment is called the Universe of the Assessment.
Let $E = \{x_1, x_2\text{…….}, x_n\}$ be a finite discrete universe of attributes x_i, $i = 1, 2, \ldots, n$.

C. Definition 2.3 Mean Fuzzy Set of IFS [3]

Let E be a universe and X be an IFS of E. The mean fuzzy set of IFS X is a fuzzy set m of E given by the membership function such that

$$\mu(x)_m = \frac{\mu_A(x) + 1 - \nu_A(x)}{2}$$

D. Definition 2.3 Total Environmental Impact Value (TEIV)

Let m be a mean fuzzy set of a finite set X and for each element $x \in X$, there is an associated weight $W_x \in R+$ (set of all non-negative real numbers) , then the TEIV of the fuzzy set m is the non-negative number given by $TEIV(m) = \sum [\mu(x_i)_m.W_i]$, $i = 1, 2, 3, \ldots\ldots, n$. Naturally maximum TEIV will indicate logically more degree of certainty in support of greenness of the building.

E. Definition 2.4 Fuzzy Decision (FD)

As a human being all decision makers or engineers hesitate more or less on every evaluation activity due to limitation of knowledge or intellectual functionaries. The fuzzy decision (FD) is an appropriate tool to achieve the targeted goal where this paper cost of land and construction of a building are the two constraints to achieve the goal of the most economical green building. To understand the function of FD an algorithm is presented below.

F. Definition 2.4 Algorithm of FD

Logically in fuzzy decision, the membership value (μ) for the maximum favourable condition of a given goal or constraint is treated as 1 and for minimum it is 0.
Let us consider a group of options as O
Where, $O = \{o_1, o_2, o_3, \ldots\ldots, o_L\} = \{ o_i \}$,
for $i = 1, 2, 3, \ldots, L$
Let a fuzzy set G describing goals associated with each option (o_i) such that

$$G = \{\mu(g_1/o_1), \mu(g_2/o_2), \mu(g_3/o_3), \ldots\ldots, \mu(g_L/o_L)\}$$
$$= \{\mathbf{\mu(g_i/o_i)}\}, \text{ for } i = 1, 2, 3, \ldots, L$$

Now if the two fuzzy sets C_1 and C_2 describing two constraints associated with each option (o_i) such that

$$C_1 = \{ \mu_1(c_1/o_1), \mu_1(c_2/o_2), \mu_1(c_3/o_3), \ldots\ldots, \mu_1(c_L/o_L) \}$$
$$= \{\mathbf{\mu_1(c_i/o_i)}\}, \text{ for } i = 1, 2, 3, \ldots, L$$

And $C_2 = \{ \mu_2(c_1/o_1), \mu_2(c_2/o_2), \mu_2(c_3/o_3), \ldots., \mu_2(c_L/o_L) \}$
$= \{\mathbf{\mu_2(c_i/o_i)}\}$, for $i = 1, 2, 3, \ldots, L$
Then the Fuzzy Decision (FD) will be given by
$FD = \text{Max } \{D(o_i)\}$,
Where $D(O_i) = [\text{ sub set-G} \cap \text{ sub set-}C_1 \cap \text{ subset-}C_2]$
$= \text{Min } \{\mu(g_i/o_i), \mu_1(c_i/o_i), \mu_2(c_i/o_i) \}$
Now to validate the fuzzy model, a case study is presented below.

III. CASE STUDY

Consider a project "SELECTION OF MOST ECONOMICAL GREEN BUILDING OUT OF n-ALTERNATIVES" of same plan, design and orientation to be constructed in different cities of a country. To do this project, we use IFS tool for assessment of TEIV of each green building and thereafter use FD-model to decide which one is the most economical out of all. For the sake of simplicity and to understand the methodology, we consider only ten green buildings (GB) of ten different cities (CT), and for evaluation of fuzzy data the following ten favourable attributes have been considered for assessment:-

x_1 = high reflective roof coating
x_2 = good indoor environmental quality
x_3 = good storm water management
x_4 = good approach
x_5 = good sanitary fittings
x_6 = good drainage system
x_7 = good water quality and quantity
x_8 = adequate natural light deep into the interior spaces
x_9 = low energy consumption
x_{10} = good habitable outdoor space

Suppose the data collected from five expert's perception for

many constraints are clubbed with all possible decisions. In This data leads to the fuzzy set X of the universe E, where

E = {x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10} },

Now the job is to assign the values of these attributes for each green building (GB) by collecting views and perception from a good number of experts individually and independently. Here we select five experts from each city for assessment of TEIV of each building separately.

the green building-GB_1 of city-CT_1 is presented in Table-1. The average perception of all experts is obviously the element of fuzzy subset $GB_1 \in E$, where, **GB_1** = {(x_1, .80, .15), (x_2, .85, .05), (x_3, .70, .10), (x_4, .70, .20), (x_5, .65, .25), (x_6, .60, .20), (x_7, .50,.30), (x_8, .40, .50), (x_9, .55, .10) , (x_{10}, .85, .10)}.

(Table-1: Average Perceptions of Five Experts)

Expert → / Attributes ↓	Expert-1	Expert-2	Expert-3	Expert-4	Expert-5	Average Perception
x_1	(x_1,.75, .20)	(x_1,.90, .05)	(x_1,.75, .20)	(x_1,.85, .10)	(x_1,.75, .20)	(x_1,.80, .15)
x_2	(x_2,.85,.10)	(x_2,.88,.07)	(x_2,.82,.05)	(x_2,.87,.01)	(x_2,.83,.02)	(x_2, .85, .05)
x_3	(x_3,.65,.05)	(x_3,.70,.15)	(x_3,.80,.11)	(x_3,.75,.09)	(x_3,.60,.10)	(x_3, .70, .10)
x_4	(x_4,.60,.20)	(x_4,.80,.15)	(x_4,.64,.25)	(x_4,.76,.15)	(x_4,.70,.25)	(x_4, .70, .20)
x_5	(x_5,.66,.24)	(x_5,.71,.21)	(x_5,.67,.25)	(x_5,.62,.26)	(x_5,.59,.29)	(x_5, .65, .25)
x_6	(x_6,.63,.16)	(x_6,.54,.24)	(x_6,.60,.30)	(x_6,.57,.18)	(x_6,.66,.12)	(x_6, .60, .20)
x_7	(x_7,.50,.26)	(x_7,.46,.34)	(x_7,.53,.31)	(x_7,.50,.32)	(x_7,.51,.27)	(x_7, .50,.30)
x_8	(x_8,.41,.47)	(x_8,.40,.44)	(x_8,.39,.52)	(x_8,.39,.53)	(x_8,.41,.54)	(x_8, .40, .50)
x_9	(x_9,.52,.08)	(x_9,.58,.12)	(x_9,.51,.07)	(x_9,.57,.11)	(x_9,.57,.12)	(x_9, .55, .10)
x_{10}	(x_{10},.79,.16)	(x_{10},.80,.15)	(x_{10},.88,.07)	(x_{10},.86,.09)	(x_{10},.92,.03)	(x_{10}, .85,.10)

Now suppose weight of each attributes of set GB_1 are prefixed by the five experts of city-CT_1 like as for $x_1 = 80$, for $x_2 = 90$, for $x_3 = 55$, for $x_4 = 40$, for $x_5 = 15$, for $x_6 = 35$, for $x_7 = 50$, for $x_8 = 70$, for $x_9 = 80$, and for $x_{10} = 95$ respectively. Then the overall TEIV of GB_1 of city-CT_1, calculated by the definition-2.4 is given below (Table 2).

(Table 2 : Overall TEIV of GB_1)

Attribute name	in support of truthness $\mu(x)$	in support of falseness $\nu(x)$	indeterministic part $\pi(x_i)$	weight of the attribute W_x	**TEIV for GB1**
x_1	.80	.15	.05	80	
x_2	.85	.05	.10	90	
x_3	.70	.10	.20	55	
x_4	.70	.20	.10	40	
x_5	.65	.25	.10	15	**475**
x_6	.60	.20	.20	35	
x_7	.50	.30	.20	50	
x_8	.40	.50	.10	70	
x_9	.55	.10	.35	80	
x_{10}	.85	.10	.05	95	

Similarly, the TEIVs of other green buildings of rest nine cities are calculated based on the data collected from the individual five experts of nine cities independently which are as follows :

GB_2 = {(x_1, .75, .20), (x_2 , .80, .15), (x_3, .65, .20), (x_4 , .80,.10), (x_5, .75, 15), (x_6, .65 .20), (x_7, .55,.15), (x_8 , .55, .20), (x_9, .70, .20) , (x_{10} , .80, .15)}, and **TEIV = 386**

GB_3 = {(x_1, .70, .25), (x_2 , .70, .15), (x_3, .80,. 15), (x_4, .55, .30), (x_5, .70, .10), (x_6, .75, .10), (x_7, .80, .10), (x_8 , .50, .25), (x_9, .60, .30) , (x_{10} , .75, .10)} and **TEIV = 448**

GB_4 = {(x_1, .60, .30), (x_2 , .75, .20), (x_3, .60 ,.20), (x_4 , .65, .10), (x_5, .60, .25), (x_6, .65 .25), (x_7, .45, .50), (x_8 , .30, .60), (x_9, .50, .30) , (x_{10} , .50, .35)} and **TEIV = 357**

GB_5 = {(x_1, .65, .20), (x_2 , .80, .10), (x_3, .75, .10), (x_4 , .90, .05), (x_5, .65, .15), (x_6, .80, .15), (x_7, .60, .25), (x_8 , .50, .35), (x_9, .80, .10) , (x_{10} , .75, .10)} and **TEIV = 422**

GB_6 = {(x_1, .85, .10), (x_2 , .80, .10), (x_3, .70, .20), (x_4, .60, .20), (x_5, .65, .30), (x_6, .70 .10), (x_7, .65, .30), (x_8 , .80, .10), (x_9, .35, .55) , (x_{10} , .45, .30)} and **TEIV = 435**

GB_7 = {(x_1, .60, .30), (x_2, .75, .10), (x_3, 85, .10), (x_4 , .55, .20), (x_5, .85, .10), (x_6, .70, .20), (x_7, .50,. 40), (x_8, .60, .25), (x_9, .80, .15) , (x_{10} , .65, .15)} and **TEIV = 475**

GB_8 = (x_1, .60, .20), (x_2 , .85, .10), (x_3, .45, .45), (x_4 , .70, .25), (x_5, .75, .20), (x_6, .70, .20), (x_7, .70, .25), (x_8 , .50, .30), (x_9, .65, .25) , (x_{10} , .70, .15)} and **TEIV = 365**

GB_9 = {(x_1, .65, .30), (x_2 , 75, .15), (x_3, .70, .10), (x_4 , .65,. 25), (x_5, .70, .15), (x_6 , .60, .35), (x_7, .55, .20), (x_8 , .40, .25), (x_9, .70, .20) , (x_{10} , .60, .30)} and **TEIV = 345**

GB_{10} = {(x_1, .70, .20), (x_2, .75, .15), (x_3, .85, .10), (x_4, .90, .05), (x_5, .65, .30), (x_6, .75, .15), (x_7, .75, .10), (x_8 , .50, .35), (x_9, .70, .10) , (x_{10} , .50, .30)}and **TEIV = 323**

Now, the task is to follow the algorithm of Fuzzy Decision (FD) considering these TEIVs and cost of land and construction. Thus, finally conclude which one is the most economical Green Building (GB) out of 10 options considered. All the data for the cost of land and construction of building (Table-3) are precise data and will not deviate by the expert's perception in anyway.

(Table-3 : Individual Data of Ten GBs)

GB	TEIV	Cost of Construction (Rs. In lakhs)	Cost of Land (Rs. In lakhs)
GB_1	475	22	45
GB_2	386	34	30
GB_3	448	25	60
GB_4	357	38	90
GB_5	422	39	85
GB_6	435	45	55
GB_7	475	32	87
GB_8	365	48	69
GB_9	345	45	53
GB_{10}	323	36	90

Now for the most economical green building, the 'TEIV' will be the goal, i.e. **G** and if we consider 'Cost of Construction' and 'Cost of Land' as the two constraints, i.e. $\mathbf{C_1}$ and $\mathbf{C_2}$, then the fuzzy sets for each options will be as :

$G = \mu(g_i/GB_i) = [1.0/GB_1, 0.50/GB_2, 0.85/GB_3, 0.75/GB_4, 0.60/GB_5, 0.70/GB_6, 1.0/GB_7, 0.40/GB_8, 0.25/GB_9, 0.15/GB_{10}]$

$C_1 = \mu(C_1/GB_i) = [1.0/GB_1, 0.65/GB_2, 0.90/GB_3, 0.50/GB_4, 0.45/GB_5, 0.30/GB_6, 0.8/GB_7, 0.20/GB_8, 0.30/GB_9, 0.55/GB_{10}]$

$C_2 = \mu(C_2/GB_i) = [0.90/GB_1, 1.0/GB_2, 0.60/GB_3, 0.15/GB_4, 0.30/GB_5, 0.70/GB_6, 0.25/GB_7, 0.50/GB_8, 0.75/GB_9, 0.15/GB_{10}]$

Therefore, $D(GB_i) = \mu(g/GB_i) \cap \mu(C_1/GB_i) \cap \mu(C_2/GB_i)$

$= [0.90/GB_1, 0.50/GB_2, 0.60/GB_3, 0.15/GB_4, 0.30/GB_5, 0.30/GB_6, 0.25/GB_7, 0.20/GB_8, 0.30/GB_9, 0.15/GB_{10}]$

Then the fuzzy decision is given by

$$FD = \text{Max}\{D(GB_i)\} = 0.90/GB_1$$

Thus, results reveal that the Green Building-GB_1 in the city of CT_1 is the best economical green building out of ten alternatives of different cities.

IV. CONCLUSION

Building siting decisions have to be taken in inital planning stages of a project from considerations of Green Building concepts and economics of various options of cities available. For selection of most economical Green Building out of n-alternatives, no crisp data is available to facilitate decisions but subjective opinion of the experts has to be relied upon. The uncertainty in Experts' perception leads to a major decision making problem to an Engineer because the major three parameters i.e. TEIV, cost of land and cost of building are not the same for all buildings. Thus to minimise the uncertainty involved in the decision of expert, the composed model of IFS and Fuzzy Decision can only give more precise results than any other tool so far available to present day Engineers or Architects.

V. REFERENCES

[1] Abair, Jesse W.: " Green buildings: what it means to be 'green' and the evolution of green building laws ; Recent Developments in Land Use, Planning and Zoning ", The Urban Lawyer 40.3 (Summer 2008): 623(10)

[2] Atanassov, K.T. : "Intuitionistic Fuzzy Sets, Fuzzy Sets and S ystems" 20 (1986) 87-97.

[3] Biswas. Srijit : " Environmental Impact Assessment : An Intuitionistic Fuzzy Approach" International Journal of NIFS , Bulgaria. : Vol.4(2), pp.23-27 (1998)

[4] Hoffman, Andrew J., and Rebecca Henn. : " Overcoming the Social and Psychological Barriers to Green Building." Organization and Environment 21.4 (2008): 390-419.

[5] Jerry Yudelson : Green Building A to Z: understanding the language of green building(2007).

[6] Ries, Robert; Melissa M. Bilec, et al. : " The economic benefits of green buildings : "A Comprehensive Case Study" Engineering Economist 51.3 (Fall 2006): 259(37).

[7] Zadeh, L., Fuzzy Sets, Inform. Control., 8(1965), 338-353.



Sunita Bansal has done **B.E.(Civil Engg) and M.E.(Structural Engg)** from PEC University, Chandigarh in 1994 and M.Tech (Information Technology) from Tezpur University, Assam in 2003.Currently she is pursuing phD from Uttarakhand Technical University, Dehradun She is working as an Associate Professor in Manav Rachna International University, Faridabad. Her research interests include material modeling, Nanotechnology of concrete, Green Buildings, Neural Networks and fuzzy logic.



Dr. Srijit Biswas: He has obtained his Ph.D from Jadavpur University and M.E from Calcutta University,India. He is 'Fellow' of Institute of Engineers (India) and working as a Professor, Department of Civil Engineering, Manav Rachna International University, Faridabad. Besides this, he has more than 17 years research experiences out of 27 years in total. He is the member of editorial board of few reputed international journals and reviewers of many international and national journals. He authored a book on the area of "Fuzzy EIA' published from Germany.



Dr. S.K. Singh is a graduate in Civil Engineering from Gorakhpur University, Masters in Environmental Engineering from Banaras Hindu University and Ph. D in Environmental Engineering from BITS, Pilani. Presently working as Professor at Delhi Technological University. He has published around hundred papers in reputed national/international journals and chaired sessions in national and international seminars. He is member of several high level committees of Ministry of Environmental and forest, Ministry of Rural Development (Govt. of India), CAPART, AICTE, National Board of Accreditation, BIS and U.P.S.C. He is member of Expert Committee appointed by Punjab Govt. and TEQIP Delhi Govt.

© 2015 Faculty of Science and Technology, Suan Sunandha University

GIS Application in Urban Traffic Air Pollution Exposure Study: A Research Review

Amrit Kumar, Rajeev Kumar Mishra* and S. K. Singh

Department of Environmental Engineering,
Delhi Technological University, Delhi-110042, India

Abstract: This paper reports a preliminary study of the evaluation and forecast of transport-related air pollution dispersion in urban areas with the help of Geographic Information System (GIS) platform and a simulative system with graphical interface. The urban population growth, economic development, energy consumption, growing transportation demand and living standards play major role in pollution exposure in atmosphere. A lot of research has already been done to investigate the functional relationship between air quality and air pollution from transport. This study is an effort to develop a more flexible framework of model to find the exposure of the air pollution in the atmosphere. This review article describes the development of framework of different GIS inputs that help to find the exposure of vehicular pollution in megacity.

Keywords: Dispersion, Pollution Exposure, GIS Application, Urban Air Quality, Vehicular Emission.

Introduction

In current scenario the air quality remains one of the major environmental issues in modern society (Tiwary & Colls, 2010). Quality of air affects the entire human race as well as plant and animal populations on the earth. The population growth, economic development, growing transportation demand as well as living standards contribute major role to polluting the surrounding area and atmosphere (Onursal & Gautam, 1997). Urban areas produce complex gases and particulates whose characteristics depend upon a wide range of factors such as: population density, energy consumption, industrial processes, and modes of transportation and usage, which affects public health, damages agriculture, weather, and climate (Pant & Roy, 2012). At present, urban outdoor air pollution causes an estimated 1.3 million deaths per year worldwide, according to the World Health Organization (WHO, 2012). The annual growth rate of the vehicle population in India is around 10% during the previous year (Nagpure *et al.,* 2013). The GIS helps to monitor the various temporal scales and thus can reduce the cost and time of field surveys, which would play a crucial role in conservation of environment. The objective of this paper was to find out the exposure of vehicular pollution to society with the application of GIS.

Emission of toxic pollutants from motor vehicles

Motor vehicles produce more air pollutants than any other single human activity. Motor vehicle emissions from highways might be viewed as a mobile line source with an emission rate per unit length of highway. In a regional scale model, the consolidated emission from an entire urban area might be viewed as an area source with an emissions rate per unit area. In city centers and congested streets, traffic is responsible for 80% to 90% of these pollutants and this situation is particularly severe in cities of developing countries (Whitelegg & Haq, 2003; Batterman *et al*., 2014). Emissions from motor vehicles with spark ignition engines are from the exhaust, engine crankcase, and fuel system. Carbon dioxide (CO_2) and water vapor (H_2O), the main products of combustion, are emitted in vehicle exhaust (Onursal & Gautam, 1997). The major pollutants emitted from gasoline-fueled vehicles are carbon monoxide (CO), Hydrocarbons (HCs), Oxides of Nitrogen (NO_x) and lead (Pb). In most industrialized countries, people who live in urban areas tend to be more affected by allergic respiratory diseases than those who live in rural areas. Road traffic is the main contributor to air pollution in most urban areas and there is evidence that living in the vicinity of high

*Corresponding Author: rajeevmishraiitr@gmail.com

traffic roads is associated with impaired respiratory health. Transportation of freight and passengers, especially in Delhi, is primarily dependent on road traffic. As a consequence of the high traffic density in Delhi, the average speed within the city is relatively low (Gurjar *et al.,* 2008). Large quantities of CO_2, CO, HC, NO_x and particulate matter (PM) are emitted regularly from motor vehicles. The mobile air toxins such as benzene, formaldehyde, acetaldehyde, and lead, along with secondary aerosols can cause adverse human health impacts (Amato *et al*., 2010).

Factors affecting vehicular emission

Real world vehicular emissions are highly variable. Several factors account for the variability in emissions in different vehicles and the amount of environmental damage they can cause. However, due to relatively higher average temperatures, poor fuel quality, poor vehicle maintenance culture, and a high proportion of old vehicles, the level of emissions from mobile sources are usually high. Figure 1 shows the various factors that are major cause for vehicular emissions in the transport sector.

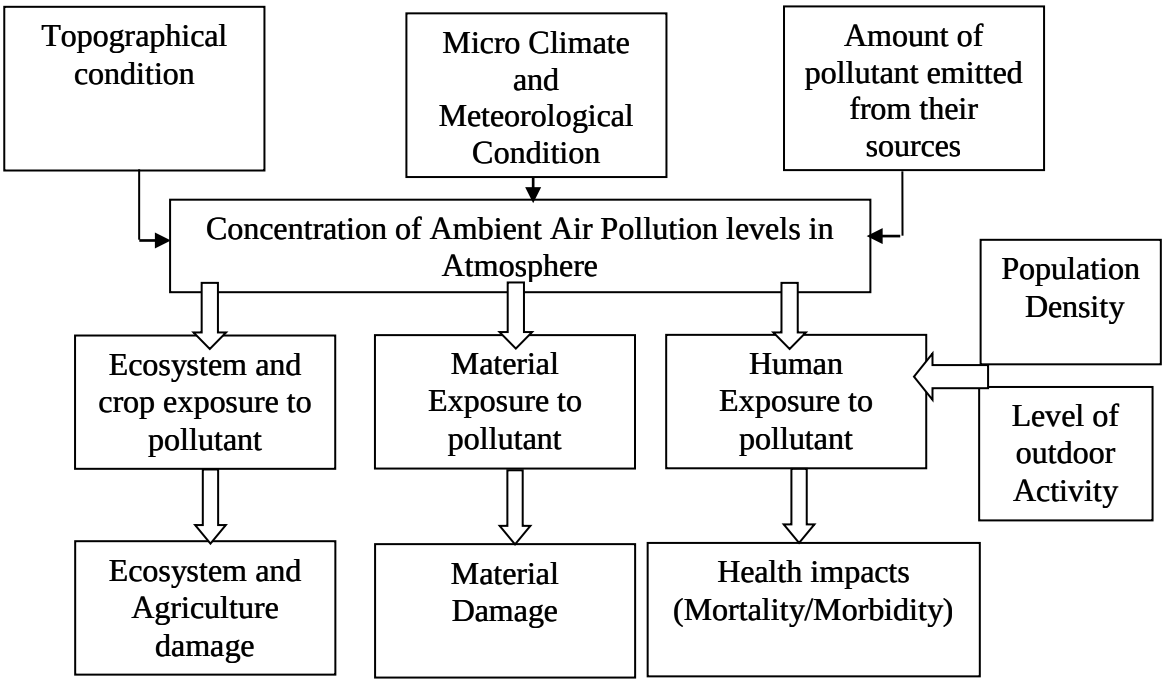


Figure 1. Factors influencing the environmental damage due to release of air pollutants.

Worldwide scenario of air pollution

Worldwide road transport vehicles like cars, trucks, and buses play a significant role in air pollution. In the European Union road vehicles are the largest sources of CO, NO_x and non-methane hydrocarbons. The recent trends indicate that road vehicles are responsible for 74% of NO_X and 94% of black smoke emissions in London (USEPA, 2010). According to the United States Environmental Protection Agency (US-EPA, 2010), diesel vehicles account for 32% and 87% of total emissions for these two pollutants respectively. Likewise, due to rapid vehicle increases, vehicle emissions have now become a major environmental concern in China (Huo *et al.*, 2011). In Indonesia, 70% of air pollution is due to transportation activity and 30% comes from industrial activity (Sutanto *et al.,* 2005). The WHO reported that the two thirds of death occur-red due to the outdoor air pollution in Asia. In 2010, particulate air pollution in Asia led to over 2.1 million premature deaths and 52 million years of healthy life lost. The outdoor air contributes to 1.2 million deaths in East Asia and 712,000 deaths in South Asia.

The World Health Organization (WHO) 2014 reported that the urban air quality database covers 1600 cities across 91 countries which are facing air pollution health risks. In April 2014, the WHO reported that outdoor air pollution was responsible for the deaths of 3.7 million people under the age of 60 in 2012. The numbers of automobiles increased at a rapid rate in mega-

cities during last decades in India. Delhi is regarded as one of the most polluted city in the world. Annual growth rates of motor vehicles are continuously increasing due to rapid motorization. The most disturbing factor is the rising density of the vehicles in urban area. The rising density of vehicles is responsible for the emissions of different vehicular pollutants like CO, NO_X, SO_X, PM_{10} and $PM_{2.5}$.

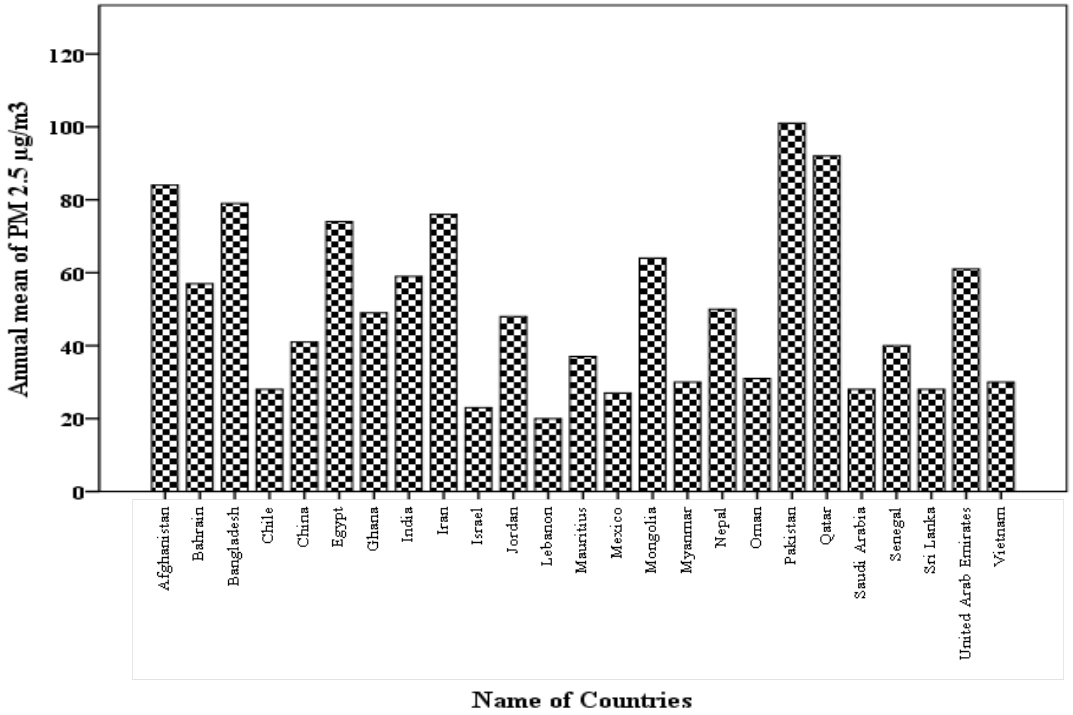


Figure 2. Annual mean of $PM_{2.5}$ µg/m^3 in various countries (Source: WHO Report, 2014).

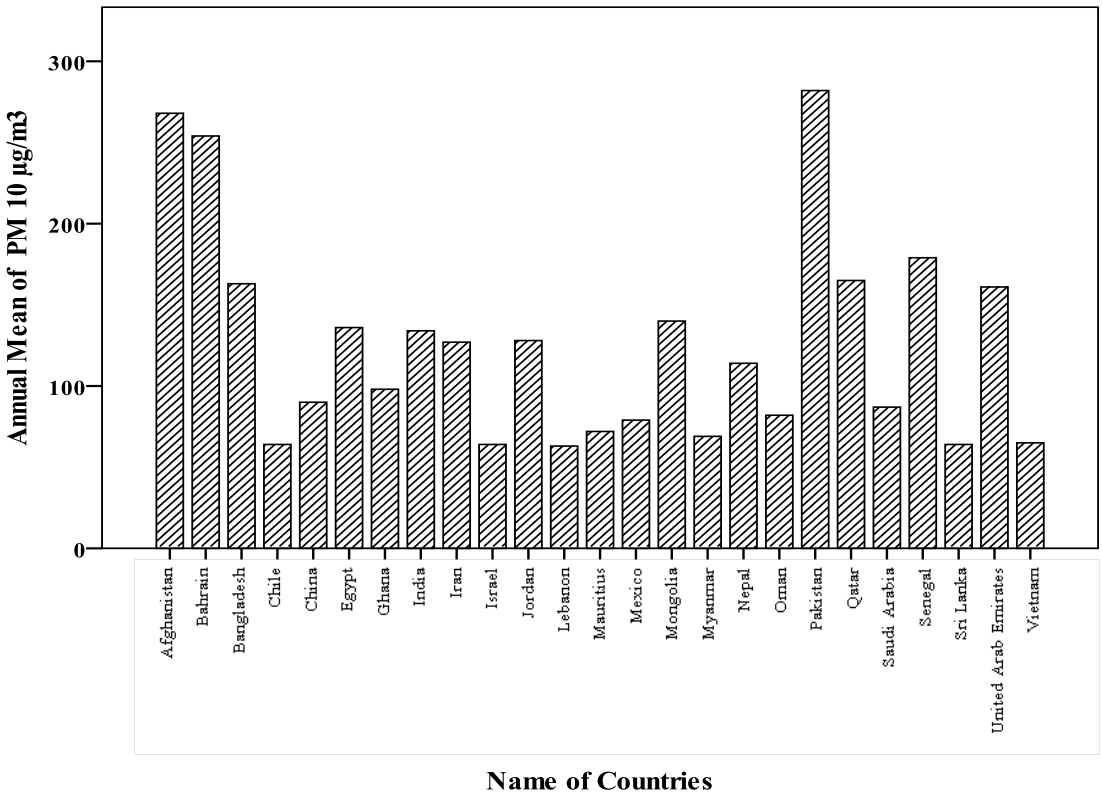


Figure 3. Annual mean of PM_{10} µg/m^3 in various countries (Source: WHO Report, 2014).

Figure 2 represents annual mean concentration of particulate matter of size 2.5µm. Various countries that exceeded the National Ambient Air Quality Standards (NAAQS) were Pakistan, Qatar, Afghanistan, Bangladesh, Iran, Mongolia, United Arab, India, Bahrain and Nepal. The lowest concentrations were observed in Lebanon. The people are also suffering due to the risk of serious health problems in these countries.

PM_{10} have higher annual mean concentrations *i.e*. more than 100µg/m^3. The human population is more exposed to PM_{10} in the respective countries. The figure 3 depicts the annual mean concentrations of PM_{10} in different countries. Some countries that exceeded the NAAQS are Pakistan> Afghanistan> Bahrain> Senegal>Qatar >Bangladesh> United Arab Emirates> Mongolia >Egypt >India >Jordan> Iran >Nepal. The lowest concentration of particulate pollutants was found in Lebanon.

Trends of air pollution in India

The outdoor air pollution has been listed as one of the top ten killers in the world, with 65 percent of the air pollution deaths occurring in Asia and close to quarter of this in India. The Central Pollution Control Board (CPCB) conducted a study in six major cities i.e. Delhi, Kanpur, Bangaluru, Pune, Chennai and Mumbai and found more than a 30% contribution of these cities in ambient air quality. The major cities like Delhi, Mumbai, Chennai, Kolkata, Bengaluru, Hyderabad and Ahmedabad, as well as secondary cities with populations of more than 2 million like Pune, Surat, Indore, Bhopal, Nagpur, Jaipur, Varanasi, Nagpur, Agra, Guwahati, Patna, Kanpur, Panaji, Trivandrum and Cochin are growing both geographically as well as in population, putting pressure on local infrastructure and life style. The increase in traffic along with inadequate infrastructure facilities is responsible for the emissions of air pollution and greenhouse gas emission (GHGs) emissions. Vehicular pollution is a major contributor towards the increase of air pollution in urban cities with respect to domestic and industrial sources, according to the Central Pollution Control Board (CPCB 2012). The following chart shows the contribution of different sectors towards the ambient air quality (Figure 4).

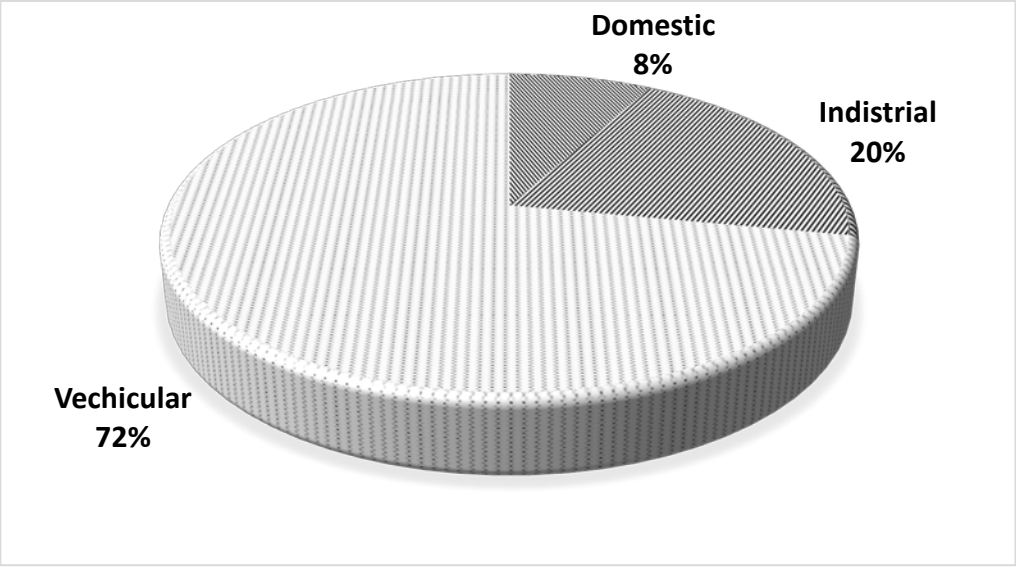


Figure 4. Contribution of various sectors to ambient air quality in urban city (Source: CPCB, 2012).

In addition to this, the annual mean of PM_{10} and $PM_{2.5}$ in different Indian cities are presented in figure 5 and 6. In figure 5, the highest concentration of $PM_{2.5}$ was found in Delhi; whereas, the highest concentration of PM_{10} was observed in Gwalior (Figure 6). Warangal city was found in much better condition in comparison to other Indian cities due to far less concentrations of $PM_{2.5}$ and PM_{10} (Figure 5 and Figure 6).

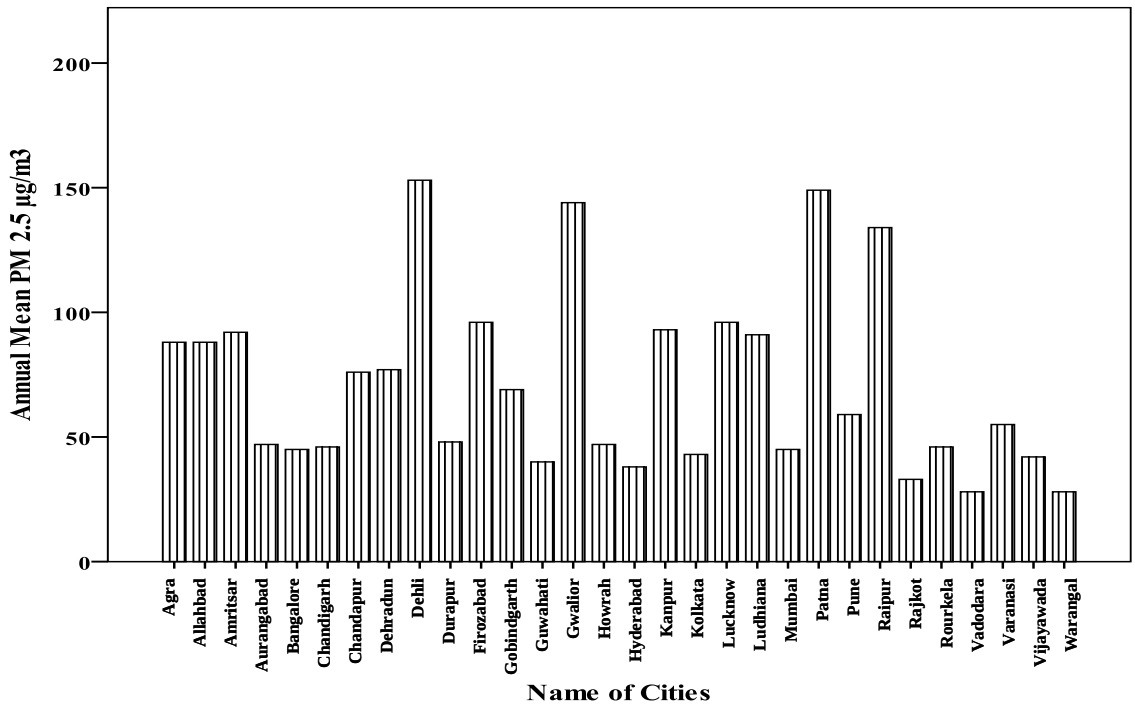


Figure 5. Annual mean of $PM_{2.5}$ µg/m^3 in various Indian cities.

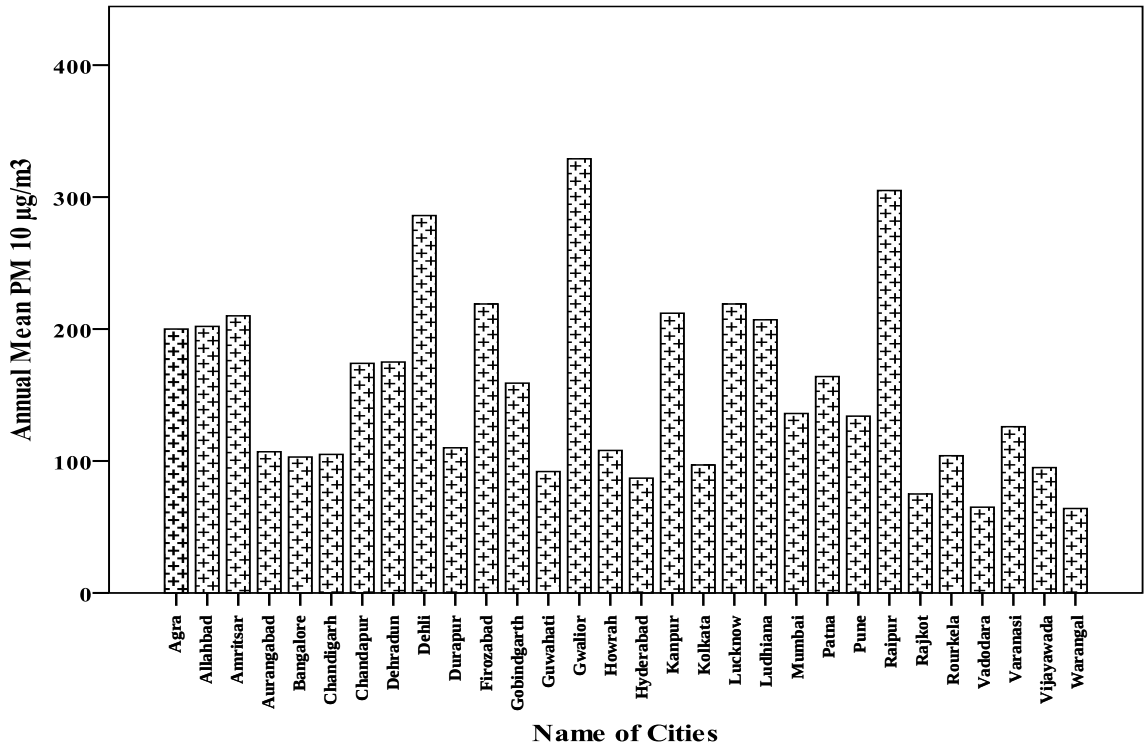


Figure 6. Annual mean of PM_{10} µg/m^3 in various Indian cities **(**Source: WHO Report, 2014).

The World Health Organization (WHO) estimates that air pollution contributes to the deaths of approximately 800,000 people annually (WHO, 2012). Different studies of megacities on different spatial scales used various modeling tools to understand their local-to-regional-to-

global impacts and implications (Guttikunda *et al*., 2001; Gurjar & Lelieveld, 2005). Outdoor air pollution causes approximately 1.3 million deaths every year worldwide (WHO, 2011). Around 40 million people within the 115 largest European cities are still exposed to air quality exceeding the European Union (EU) limit values for at least one pollutant (WHO, 2012). In megacities, people travel everyday using different types of vehicles. The daily increase in the number of people in megacities results in increased traffic volume. Every person uses vehicles according to their life style. The different types of vehicles in urban cities are buses, autos, light commercial vehicles, heavy commercial vehicles, Car/Jeep and 2-wheelers are playing a major role in the air pollution in mega cities. The buses travel more in a day, they emit more pollution and two-wheelers emit less pollution. According to a local survey, 30% of Delhi's population is suffering from respiratory disorders due to air pollution. The respiratory disorders increased 12 times as compared to all other states of India. Due to that, Delhi is characterized as the "asthma capital" of India (Guttikunda *et al.*, 2012; Das *et al.*, 2004). Increases in the death due to air pollution has made air pollution the fifth leading cause of death in India, with 620,000 premature deaths. This is up from 100,000 in 2000, a six-fold increase (WHO, 2014). Below, figure 5 and figure 6 show the annual mean of particulate matter in various cities in India.

GIS application in air quality mapping:

A geographic information system (GIS) is a computer-based tool for mapping and analyzing geographic phenomenon that exist and occur on earth. Yan *et al.* (2009), Rytkonen *et al.* (2004) and Hammond *et al*. (2011) utilized GIS to study the community level pollutant concentrations and associated health risks. They explored Toxic Assessment, Air Quality System, and National Emission Inventory for identifying risks of exposure to air pollutants at the community level. GIS basically integrates baseline data and unique visualization of geographic analysis on the maps. It is very useful to acquire the pollutant spatial distribution, variations and characteristics information because of its flexible spatial data management and effective spatial data analysis methods. It is an innovative and important component of many projects for public health and epidemiology studies. GIS may also involve more sophisticated spatial analysis of disease occurrence and contributing environmental factors. Scoggins *et al*. (2004) used model concentrations were converted from point base grid coverage to polygon grid coverage using GIS. Polygon grid coverage concentrations were then converted to census area unit concentration. It creates very informative graphical presentation and is easy to understand. Sahzabi *et al*. (2011) also conducted a study to predict CO_2 dispersion and optimization of the dispersion model. GIS is used for visualizing model results and better analyses of CO_2 concentration status at a particular geographical area. Morality rates were calculated using the number of people exposed to pollution at census area unit level. It plays a vital role in the planning of air quality management in megacities. In present scenario the GIS techniques are used for exposure air quality assessment (Nuckols *et al*., 2004). It simulates traffic, emission, and dispersion patterns of the pollution. The GIS and modeling characteristics are presently used for air quality assessment. The various sub layers are represented different in figure 8, including topographical, land use data, transport network infrastructure, socioeconomic, demographic data, traffic data, different air pollutants, and environment impact parameters. The analysis of transport planning and transport systems management can also be done through GIS technology. The GIS maps show a combination of data layers as presented in Figure 8. The GIS database is systematic on the basis of geographical base line data, which consists of inputs of attributes in the GIS software. The road corridor network map of the study area is digitized with GIS software. These linked maps are integrated with a number of various topographic morphological attributes of the study area. In particular, the whole territorial database, which has been set up within GIS, is designed according to the details of inputs. The use of GIS as a database integrator for a transport study area and schematic diagram represents a set of individual databases for the study area, which comprise a mixture of spatial, numerical, and perhaps textual data. The GIS is a very dynamic technology capable of integrating large quantities of geographic spatial data. The GIS database is provided by road networks corridor data, traffic demand characteristics, driving cycles and fleet composition. GIS integrating models are aimed at reproducing traffic behavior, emission and dispersion scenarios (Elliott & Wartenberg, 1997).

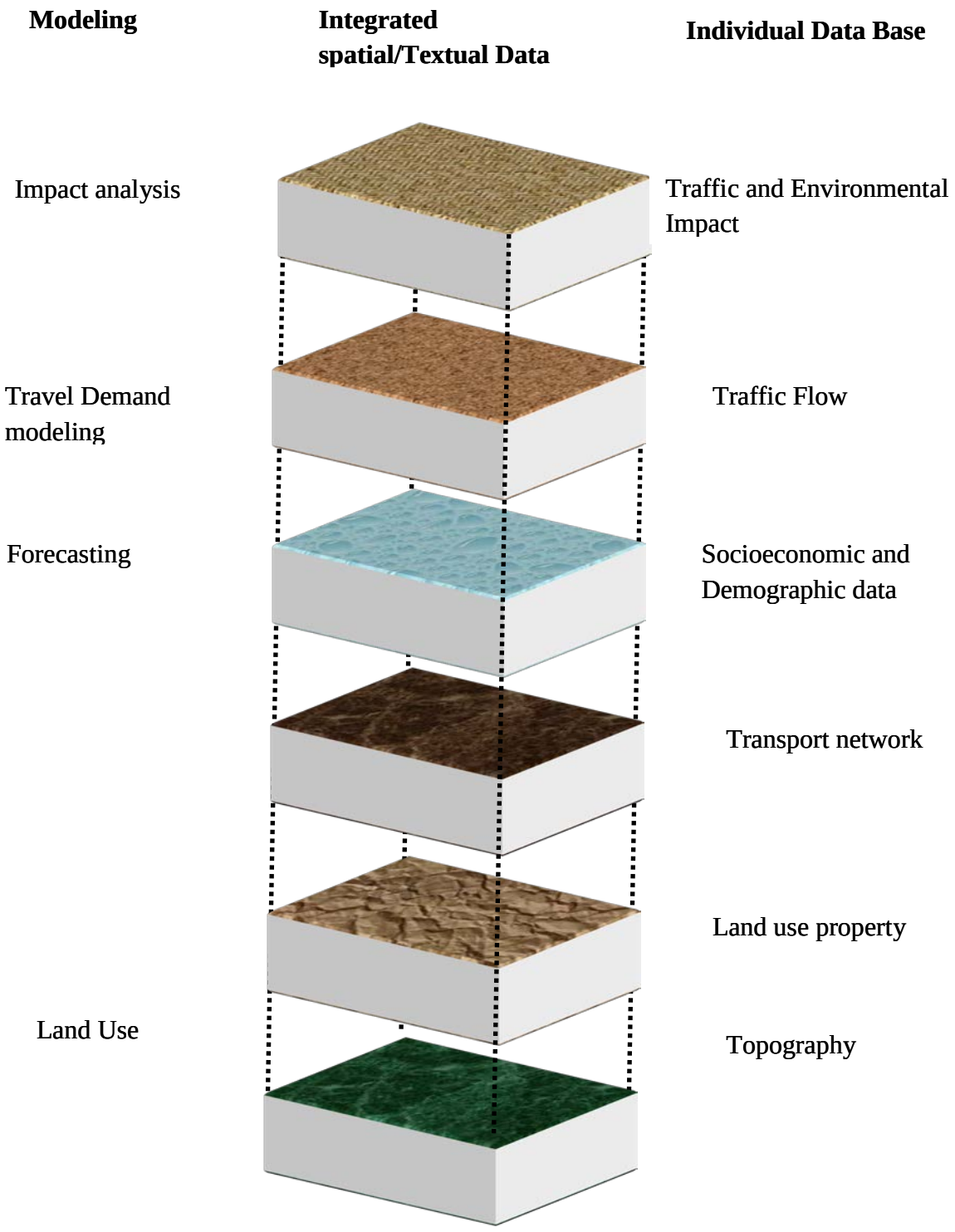


Figure 8. Layout of superposition of data layers in GIS.

Research Methodology

The principle of Geographical Information System (GIS) is basically superimposed of all spatial data and attributes of the study area. Figure 9 shows the different sub spatial data of thematic map data, transportation network data and population data. Each road link contains attribute data, which is required by the air pollution simulation model. Figure 8 also shows a combination of various input parameters, which help to develop the various models. All these models are integrated step by step with GIS and help the spatial mapping of the air pollution. The information is imported to the GIS software action to spatial datasets. The result shows the overlay to display, the intersection of sub systems as a geographical map. The base line data was collected from the different air pollution monitoring stations. The other data like meteorological parameters, different types of pollutants and their concentrations are necessary for the GIS integration for air pollution assessment.

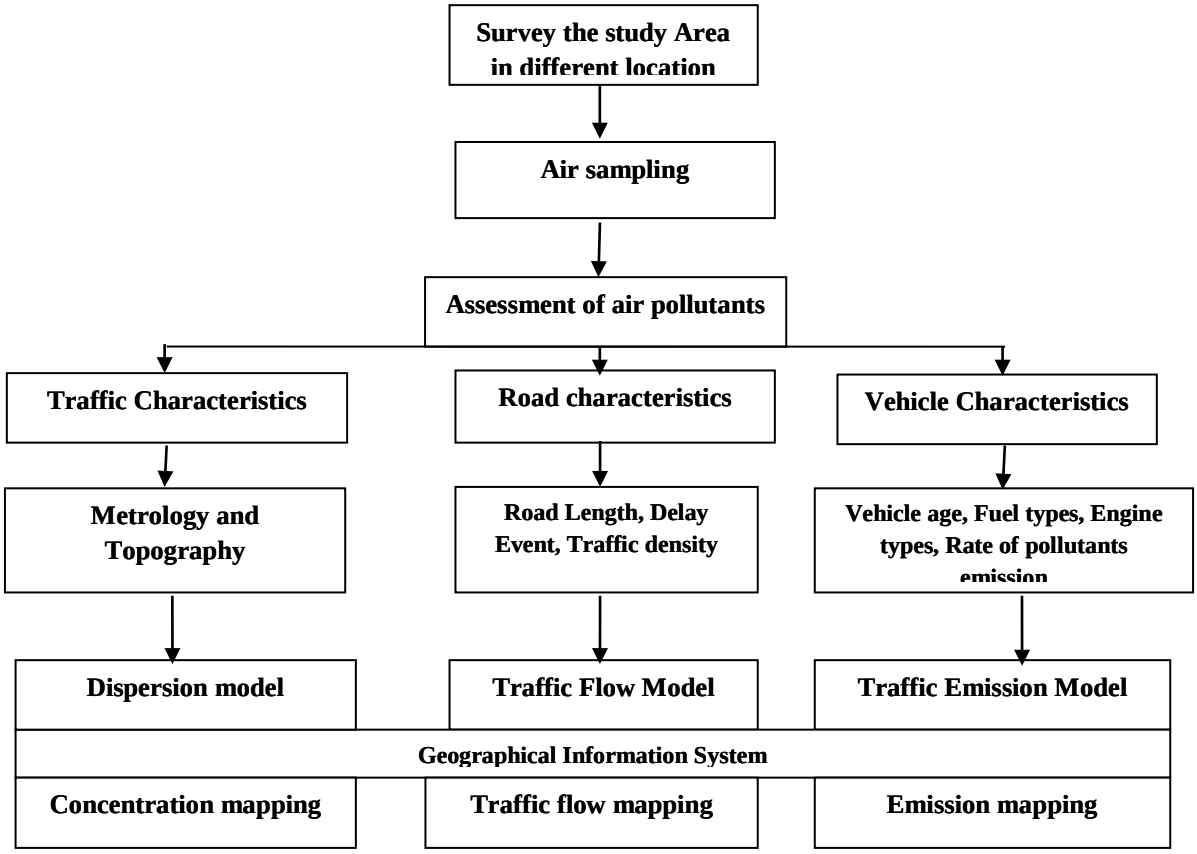


Figure 9. Air pollution exposure modeling with integration of GIS.

The base line monitoring data from pollution monitoring station include pollution concentration of various pollutants. Pollutant concentration is measured in grid zone of the study area like 2KM X 2 KM grid plot. It could be monitored more accurately and ideally using grid-based methods such as Discovery Net Project (Richards *et al.*, 2006). Spatial interpolation could be implemented in many types GIS software, such as Geomatica10, ArcGIS10.2. Exploratory Spatial Data Analysis (ESPA) is carried out according to the characteristics of the data before interpolating in order to choose the suitable interpolation methods.

Figure 10 show the mapping of the $PM_{2.5}$ and PM_{10} in the Delhi NCR. The distribution of vehicle exhaust and the brick kilns. The dark brown color shows the maximum concentration of the particulate matter. In the study, South and Central Delhi was found to be most affected zone due to particulate matter pollution.

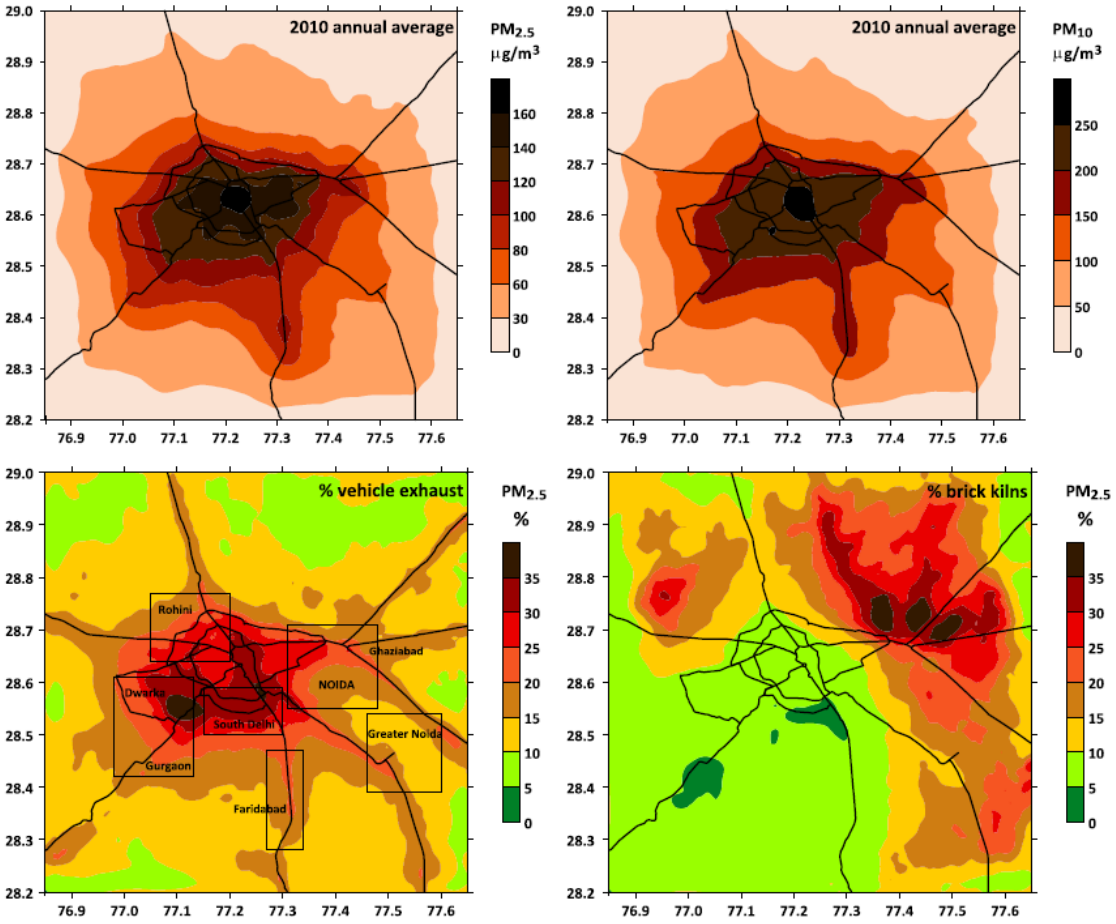


Figure 10. Annual average percentage of $PM_{2.5}$ and PM_{10} form the vehicle exhaust and brick kilns emission for year 2010 in Delhi, India (Source: Guttikunda *et al.*, 2013).

Figure 10 show the mapping of the $PM_{2.5}$ and PM_{10} in the Delhi NCR. The distribution of vehicle exhaust and the brick kilns. The dark brown color showed the maximum concentration of the particulate matter. In the study, south and central Delhi is found most affected zone due to particulate matter pollution.

Emission and dispersion models

Air quality models are used as mathematical, numerical techniques to simulate the physical, chemical processes that affect air pollutants in the environment. The different inputs like meteorological parameters, source information and emission rates of pollutants are used to characterize the primary pollutants emitted directly into the ambient air. All these models are important for the assessment and management of air quality of environment. It is very useful to identify source contributions to air quality problems and to assist in the design of effective strategies to reduce harmful air pollutants. The air quality models can also be used to predict future pollutant concentrations from multiple sources (US-EPA 2012).

The various emissions of pollutants and dispersion models simulate the pollutants. This traffic related pollutants are major source of air pollution in the urban traffic corridor. The emission estimation of various pollutants is found out with the help of emission model. The emissions of the entire vehicle fleet are calculated on the basis of emission coefficients, traffic volume, and road segments (Equation 1).

$$E_i = L \times C_i \qquad (1)$$

Where E denotes the emissions of pollutant i, L denotes the vehicle kilometers traveled, and C denotes the emission coefficient of pollutant i. Dispersion models simulate the dispersion of pollutants after its emission. Line source Gaussian plume dispersion model is originally used to

predict the impacts of carbon monoxide near different transport corridor. The Gaussian solution for linear sources which is based on is based the principle of overlaying various component, which assume the concentrations of emitted substances at the receiving equal to the sum of dispersions of all infinite points sources which is together in line source.

$$C(x,y,z) = \frac{Q}{2\pi u \sigma_y \sigma_z} exp\left(-\frac{y^2}{2\sigma^2{}_y}\right)\left[exp\left(-\frac{(z+H)^2}{2\sigma^2{}_z}\right) + exp\left(-\frac{(z-H)^2}{2\sigma^2{}_z}\right)\right] \quad (2)$$

Source of emission at ground level,

$$C(x,0,0) = \frac{Q}{2\pi u \sigma_y \sigma_z} \quad (3)$$

Where
C(x,y,z) is the concentration unit (gram/meter^3) at some point in space with coordinates x,y,z
Q= the emission rate of the pollution source (in gram/second)
u= the average wind speed in (meter/second)
σ_y =the standard deviation of the plume in the y direction (meter)
σ_y=the standard deviation of the plume in the z direction (meter)

Emission and dispersion models can be carried out using integration of GIS and environmental models. Emission models are used for the various transport corridor which gives traffic volume, gridding zone of the study area with suitable cell-size raster, finding the cells that intersect with road segments. The output raster shows the emission rate from the source area. The 2 KM X 2 KM divided as girding in the study area with suitable cell raster that calculating each cells value second equation using second equation. The obtained raster indicates the spatial distribution of pollution concentration.

Land use regression

Land use regression (LUR) combines monitoring of air pollution at a small number of locations and development of stochastic models using predictor variables usually obtained through GIS software. The regression model is then applied to a large number of locations in the study area. The input data include traffic characteristics of different land use area, physical geography. (David *et al*., 1997; Gerard *et al*., 2008).The climate variables such as wind speed direction and wind speed has significant effect on pollution dispersion distance. The climate variables are neglected when using LUR; it will bring an error to value of the predictor variables that are computed from the circular zones around each monitoring site. Gulliver *et al.* (2011, 2005) evaluated GIS-based dispersion model Space Time Exposure Modeling System-Air pollution (STEM-Air) to simulate traffic related air pollution. Behera *et al*. (2011); Mion *et al.* (2002) developed an emission inventory for all sources responsible for PM_{10} pollution in Kanpur, India, with the then used ISCST3 dispersion model to simulate PM_{10} concentrations to identify pollution hotspots in the city and to evaluate the contribution of each source in PM_{10} pollution level. Elbir *et al*. (2004) developed a decision support system for air quality management in big Turkish cities. The developed GIS system was based on the CALPUFF dispersion model and related databases to calculate emissions from various sources. Pollutant concentrations were predicted using the CALPUFF dispersion model and results were compared with observed air quality data from actual monitoring stations. After model calibration, GIS was used for air quality mapping and scenario analysis.

As shown in figure 9, the GIS is integrated with different sub models, which aimed for stimulating each sub models that process involved in the entire air pollution phenomenon. These are traffic emission estimations, dispersion models all linked together. The survey input of all geographical information which is important for the study area. Road traffic is simulated through a deterministic model designed for solving the equilibrium auto assignment problem with capacity constraints. The emission model has been developed on the basis of input data and it is able to calculate emission levels of some typical traffic-related pollutants such as carbon monoxide (CO), hydrocarbons (HC), and nitrogen oxides (NO_x). Emission factors refer to a

vehicle grouping based on the knowledge of vehicle model year, fuel type, gross weight, and power characteristics. All these input parameters are very useful for the assessments of air quality as well as for the developed model.

Exposure analysis application

Exposure assessment describes the damages to the human health, plant ecosystems and sensitive areas like hospitals, schools, residential area. Vehicular pollution exposure modeling on human is very spatially complex and temporally dynamic process. The simulation is necessary for simplification and plays a key role in implement exposure modeling (Borrego *et al*., 2006; Gulliver *et al*., 2005). GIS provides an effective way to implement exposure modeling. Exposures are estimated for each location, at each time interval, by cross-reference to the pollution concentration raster for that time period as shown in equation 4.

$$E_{ijk} = C_{ij} \times W_k \qquad (4)$$

Where E denotes the exposure of an individual during time i, at location j, in exposure environment k. C denotes pollutant concentration during time i at location j and W denotes the weighting factor for exposure environment k, relative to ambient. Modeling exposure uses raster-based calculation. Mapping exposure is carried out using map algebra operation. Exposure profiles are carried out extracting grid values on exposure map along the moving path using combination of several tools in Geomatica10.

Results and Discussion

This paper helps to examine the assessment of vehicular emissions with the help of advanced modern technology based study like GIS. Spatial distribution of urban air pollution concentration is very useful from a technology based assessment. The spatial patterns of pollution resulting from policy or other interventions and provide improved estimates of exposure for epidemicological studies. Delhi has recorded the highest increase per capita road length decrease and more importantly the number of vehicles per km of road has tremendously increased. This indicates the number of vehicles registering has increased fivefold in Delhi. The rapid growth of personalized transport in Indian cities gives rise to serious issue due to increment of air pollutant problems. In the current scenario, Delhi is facing the highest vehicular emission load. In recent times, Delhi NCR is rapidly modernizing and commercializing. Figure 2 shows the air pollution mostly due to vehicular emissions which contributes 72% and the industrial sector contributes 20% in the megacities of India. This shows that the road transportation has a major role in air pollution. The considerable magnitude of air pollution increases the number of people suffering from respiratory and cardiovascular diseases. That is the major issue which results in premature deaths in India. There is a need to control population growth and vehicular air pollution in the megacities. Special efforts should be made for educating the general mass and local leaders about the adverse effects of a large population and vehicular pollution through specially designed information, education and communication activities. Upgrading the quality of Indian fuel, enforcing higher emission standards and regulating traffic can reduce the pollution caused by the explosion in the number of automobiles in megacities. Those who use personal vehicles should be encouraged to use the public transport system. State transport service should set up better vehicular standards, and more research and development should be encouraged in vehicle technology. Air pollution should not be a responsibility of government alone but mass and local leaders should be encouraged to make dedicated efforts to eradicate the air pollution problems. The policy aimed at overall development should certainly include efforts to control population, private vehicles, and air pollution for better health of the present and future generations.

Conclusions

In present scenario the developing as well as developed megacities are polluting the ambient air and disturbing the natural environment system. It also increases public health issues. It is necessary to learn comprehensive processes and ways to study urban pollution. This paper represents the air quality modeling and evaluating the impact of vehicular pollution with advanced technology. In this study, the integrated modeling approach involving GIS has been used for pollution mapping of vehicular pollutants in megacities. It monitors various temporal scales, which will reduce the cost and time of field surveys. It is most suitable for the local administration in order to forecast alert pollutant levels in air quality area. It also gives the integrated

existing monitoring network measurements and estimates the sensitivity of pollution levels to such variables as traffic flow distribution and atmospheric conditions. The interpretation of satellite data is more complex as compared to in-situ observations. Satellite data sets should be of good quality, easily accessible and reliable for the assessment of air pollution. It is very useful in the model-based approach and the predictions for the epidemiology remain to be seen. The desired future development of GIS requires a switch of emphasis from data and information to knowledge of the atmospheric conditions. The results help in the developments of methods for the study of ambient air pollution and air pollution assessment. GIS has been proved a powerful tool to implement the methods involved in the framework and makes the methods more efficient and flexible. Urban air pollution research will be further developed with increasing concerns of urban planners, developers, health officials and citizens living in the city. The GIS framework is an extension of new methods for the environmental conservation of future.

References

Amato, G. D., L. Cecchi, M. D. Amato & G. Liccardi. 2010. Urban Air Pollution and Climate Change as Environmental Risk Factors of Respiratory Allergy an Update. *J. Investig. Allergol. Clin. Immunol.* 20(2): 95-102.

Batterman S., Chambliss S & Isakov V. 2014. Spatial resolution requirements for traffic-related air pollutant exposure evaluations *Atmospheric Environment* 94 (2014) 518-528

Behera S. N., M. Sharma, O. Dikshit & S. P. Shukla. 2011. GIS-Based Emission Inventory, Dispersion Modeling, and Assessment for Source Contributions of Particulate Matter in an Urban Environment. *Water Air Soil Pollution* 218: 423-436.

Borrego C., O. Tchepel & A. M. Costa. 2006. Traffic-related particulate air pollution exposure in urban areas. *Atmospheric Environment* 40: 7205-7214.

Brunsdon F, & Charlton. Geographically weighted regression: A method for exploring spatial non stationnarity. *Geographical Analysis* 28: 281–298.

Central Pollution Control Board (CPCB) Annual Report (2011-2012). 2012. Prepared and Published by PR Division. *Central Pollution Control Board.*

Das, A. & J. Parikh. 2004. Transport scenarios in two metropolitan cities in India: Dehli and Mumbai. *Energy Conversion and Management* 45: 2603-2625.

David, J., C. Briggs Sunsan & E. Paul. 1997. Mapping urban pollution using GIS: A regression-based approach. *International Journal of Geographical Information Science* 11: 699-718.

Elbir T. 2004. A GIS based decision support system for estimation, visualization and analysis of air pollution for large Turkish cities. *Atmospheric Environment* 38: 4509-4517.

Elliott P, & D. Wartenberg. 1997. Spatial epidemicology current approaches and future challenges. *Environmental Health Perspectives* 1997; 105(6): 598-605.

Faiz A, S. W. Weaver & M. P. Walsh. 1996. *Air Pollution from Motor Vehicles: Standards and Technologies for Controlling Emissions.* Washington, D.C: The World Bank.

Gerard H., B. Rob & K. de Hoogh. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution, *Atmospheric Environment* 42: 7561-7578

Gulliver J. & D. Briggs. 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Science of the Total Environment* 409: 2419–2429.

Gulliver J., & D. J. Briggs. 2005. Time-space modeling of journey-time exposure to traffic-related air pollution using GIS. *Environmental Research* 97: 10-25.

Gurjar B. R., V. J. A. Aardennea, J. Lelievelda & M. Mohan. 2004. Emission estimates and trends (1990–2000) for megacity Delhi and implications. *Atmospheric Environment* 38: 5663–5681.

Gurjar B.R., & J. Lelieveld. 2005. New Directions: Megacities and global change. *Atmospheric Environment* 39(2): 391–393.

Gurjar B.R., T. M. Butler, M. G. Lawrence, & J. Lelieveld. 2008. Evaluation of emissions and air quality in megacities. *Atmospheric Environment* 42: 1593–1606.

Guttikunda S. K. & R. G. Bhola. 2012. Role of meteorology in seasonality of air pollution in megacity Delhi, India *Environmental Monitoring and Assessment* 184(5): 3199-3211.

Guttikunda S. K., Y. H. Tang, G. R. Carmichael, G. Kurata, L. Pan, D. G. Streets, J. H. Woo, N. Thongboonchoo & A. Fried. 2001. Impacts of Asian megacity emissions on regional air

quality during spring. *J. Geophys. Res*. 110 (D20): 1-18.

Guttikunda, S. K., G. Calori. 2013. A GIS based emissions inventory at 1KM X 1KM spatial resolution for air pollution analysis in Delhi India. *Atmospheric Environment* 67: 101–111.

Hammond D., K. Conlon, T. Barzyk, T. Chahine, V. Zartarian & B. Schultz. 2011. Assessment and Application of National Environmental Databases and Mapping Tools at the Local Level to Two Community Case Studies. *Risk Analysis* 31:43-53.

Huo, H., Q. Zhang, K. He, Z. Yao, X. Wang, B. Zheng, D. G. Streets, Q. Wang, & Y. Ding. 2011. Modeling vehicle emissions in different types of Chinese cities: Importance of vehicle fleet and local features. *Environmental Pollution* 159: 2954-2960.

Min, D.L. & C. L. Yung. 2002. The application of GIS to air quality analysis in Taichung City, Taiwan, ROC, *Environmental Modeling & Software* 17: 11-19.

National Ambient Air Quality Standards (NAAQS). 2012. *Report of national ambient air quality monitoring* 35 /2011-2012.

Nagpure A.S., K. Sharma & B. R. Gurjar. 2013. Traffic induced emission estimates and trends (2000–2005) in megacity Delhi. *Urban Climate* 4: 61–73.

Nuckols J.R., M. H. Ward & L. Jarup. 2004. Using geographic information systems for exposure assessment in environmental epidemicological studies. *Environmental Health Perspective* 112: 1007–1015.

Onursal, B. & S. P. Gautnam. 1997. *Vehicle Air Pollution: Experiences from Seven Latin American Urban Centers*. Technical Paper No. 373. Washington D.C. World Bank.

Pant, P. & M. Roy. 2012. Harrison Critical Review of Receptor Modeling for Particulate Matter: A Case Study of India. *Atmospheric Environment* 49: 1-12.

Richards, M., M. Ghanem & M. Osmond. 2006. Grid-based analysis of air pollution data. *Ecological Modeling* 94: 274-

Rytkonen, M. J. 2004. Not all maps are equal GIS and special analysis in epidemiology. Int. J. Circumpolar Health 63(1): 9-24.

Sahzabi, A. Y., K. Sasaki, H. Yousefi, S. Pirasteh, & Y. Sugai. 2011. GIS aided prediction of CO_2 emission dispersion from geothermal electricity production. *Journal of Cleaner Production* 19: 1982-1993.

Scoggins A, T. Kjellstrom, G. Fisher, J. Connor, & N. Gimson. 2004. Spatial analysis of annual air pollution exposure and mortality. *Science of the Total Environment* 321:71-85.

Sutanto, S. & S. T. Erni. 2005 Study on correlation between motor vehicle emission and public health, *Eastern Asia Society for Transportation Studies* 5: 1841-1856.

Tiwary, A, J. Colls. 2010. *Air Pollution: Measurement, Modeling & Mitigation*. Routledge. Taylor & Francis Group, Abingdon, Oxon, UK. 2010. Third edition.

United States Environmental Protection Agency (USEPA). 2012. *AERMOD Evaluations Under the 1-hour NO2 and SO2 NAAQS* available at http://www.epa.gov/ttn.

United State Environmental Protection Agency (USEPA). 2010. *Air Quality Index (AQI) - A Guide to Air Quality and Your Health*. http://www.epa.gov/2010.

Vine, M. F., D. Degnan & C. Hanchette. 1997. Geographic information systems: their use in environmental epidemiologic research *Environmental Health Perspective* 105(6): 598-605

Whitelegg, J. & G. Haq. 2003. *The Global Transport Problem: Some Issues but a Different Place*, in John World Transport, Policy and Practice. London: Earth Scan Publications Limited.

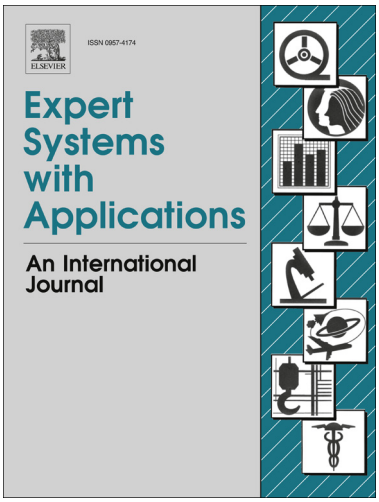
World Health Organization (WHO). 2014. *Air Quality and Health*. from: http://www.who.int/.

World Health Organization (WHO). 2011. *Air Quality and Health. Report 2011*. http://www.who.int/

World Health Organization (WHO). 2012. *Air Quality and Health*. Retrieved March 2012, from: http://www.who.int/mediacentre/factsheets/fs313/en/index.html.

Yang, K, C. Man & C. Zhen. 2009. Urban Air Pollution Study Based on GIS. *School of Geographic and Oceanographic Sciences, Nanjing University Nanjing, China IEEE* 978(1): 4244-3461.

Accepted Manuscript



Hybrid Classifier based Human Activity Recognition using the silhouette and cells

D.K. Vishwakarma, Rajiv Kapoor

PII: S0957-4174(15)00281-X
DOI: http://dx.doi.org/10.1016/j.eswa.2015.04.039
Reference: ESWA 9986

To appear in: *Expert Systems with Applications*

Please cite this article as: Vishwakarma, D.K., Kapoor, R., Hybrid Classifier based Human Activity Recognition using the silhouette and cells, *Expert Systems with Applications* (2015), doi: http://dx.doi.org/10.1016/j.eswa.2015.04.039

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

Hybrid Classifier based Human Activity Recognition using the silhouette and cells

D. K. Vishwakarma*, Rajiv Kapoor
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 110042, India.

Abstract: The aim of this paper is to present a new approach for human activity recognition in a video sequence by exploiting the key poses of the human silhouettes, and constructing a new classification model. The spatio-temporal shape variations of the human silhouettes are represented by dividing the key poses of the silhouettes into a fixed number of grids and cells, which leads to a noise free depiction. The computation of parameters of grids and cells leads to modelling of feature vectors. This computation of parameters of grids and cells is further arranged in such a manner so as to preserve the time sequence of the silhouettes. To classify, these feature vectors, a hybrid classification model is proposed based upon the comparative study of Linear Discriminant Analysis (LDA), K-Nearest Neighbour (K-NN) and Support Vector Machine (SVM) Classifier. The proposed hybrid classification model is a combination of SVM and 1-NN model and termed as 'SVM-NN'. The effectiveness of the proposed approach of activity representation and classification model is tested over three public data sets i.e. Weizmann, KTH, and Ballet Movement. The comparative analysis shows that the proposed method is superior in terms of recognition accuracy to similar state-of-the-art methods.

Key Words: Human Activity Recognition (HAR); Linear Discriminant Analysis; K-nearest Neighbour; Support Vector Machine; Hybrid Classifier.

1. Introduction

In recent years, the area of vision based Human Activity Recognition (HAR) has become an important area of research in computer vision, due to its various applications: Surveillance, Assistive health care, Content based video analysis, Interaction between people, Sports, Robotics, and Prevention of terrorist activities (Agrawal & Ryoo, 2011; Chaaraoui, Pérez, & Revuelta, 2012; Vishwakarma & Agrawal, 2012).

The task of the HAR system is to detect and analyse human activity/action in a video sequence. The reviews of previous work (Agrawal & Ryoo, 2011; Poppe, 2010;

* Corresponding Author.

ACCEPTED MANUSCRIPT

Vishwakarma & Agrawal, 2012; Weinland, Ronfard, & Boyer, 2011) reveals the challenges in vision based HAR systems. The various factors that make the task challenging are the variations in body postures, the rate of performance, lighting conditions, occlusion, view point and cluttered background. A good HAR system is capable of adapting to these variations and efficiently recognizes the human action class. The important steps involved in HAR systems are usually: a) Segmentation of foreground b) Efficient extraction and representation of feature vectors, and c) Classification. An efficient and novel solution can be proposed at any step of the work individually, or collectively for all the steps. Due to the variation in human body taxonomy and environmental conditions, every step is full of challenges and therefore, one can only provide the best solution in terms of recognition accuracy and processing speed. The Shape and Motion feature based descriptors (Agrawal & Ryoo, 2011) are two widely used methods in HAR systems. Shape based descriptor is generally represented by the silhouette of the human body and silhouettes are the heart of the action. Motion based descriptors are based on the motion of the body, and the region of interest can be extracted using optical flow, and pixel wise oriented difference between the subsequent frames. The motion based descriptors are not efficient, especially when the object in the scene is moving with variable speed.

The main contributions of this paper are twofold: Firstly, for effective representation of human activity, a texture based background subtraction approach is used for the extraction of silhouettes of the human activity from the video sequence. The key poses of the silhouettes are chosen and described by forming cells and grids to produce the descriptor. The modelling of feature vector is done through the easy computation of parameters of grids and cells. This modelling is further arranged in such a manner as to preserve the time sequence of the silhouettes. Secondly, to improve the classification accuracy of HAR system, a hybrid classification model of “SVM-NN” is constructed.

The rest of the paper is organized as follows: section 2 presents the past work carried out in the field of human activity recognition. The details of the proposed framework which comprises silhouette extraction, feature extraction, feature representation, and various classifiers are presented in section 3. Section 4 gives the details of experimental work, and the discussion of the result.

2. Related Works

ACCEPTED MANUSCRIPT

Significant amount of work has been reported in the literature for the recognition of human action and activity using video sequences and most of the HAR methods rely on the local features, global features, key points, spatial-temporal features, bags of words etc. (Agrawal & Ryoo, 2011; Chaaraoui, Pérez, & Revuelta, 2012; Poppe, 2010; Vishwakarma & Agrawal, 2012; Weinland, Ronfard, & Boyer, 2011; Ziaeefar & Bergevin, 2015). All these methods generate a set of features and then an appropriate machine learned classifier is used for the recognition of the activity. A brief review of local feature based spatio-temporal interest point (STIP) and holistic approach that incorporate both local as well as global features are discussed.

An efficient approach of spatio-temporal interest points based on local features using a temporal Gabor filter and a spatial Gaussian filter was introduced by Dollar et al. (2005). Thereafter, a number of STIPs detectors and descriptors have been proposed by several researchers (Chakraborty, Holte, Moeslund, & Gonzàlez, 2012; Everts, Gemert, & Gevers, 2014; Jargalsaikhan, Little, Direkoglu, & O'Connor, 2013; Laptev, 2005; Ryoo & Aggarwal, 2009). These local features based descriptors became popular due to their robustness against noise, illumination change, and background movements. However, these methods seemingly, are less effective for complex activity modelling (e.g. Ballet movement).

A holistic approach of human action recognition that relies on the human silhouette sequences was proposed by several researchers (Bobick & Davis 2001; Chaaraoui & Revuelta, 2014; Gorelick, Blank, Shechtman, Irani, & Basri, 2007; Olivieri, Conde, & Sobrino, 2012; Eweiwi, Cheema, Thurau, & Bauckhage, 2011; Weinland, Boyer, & Ronfard, 2007; Wu & Shao, 2013). In silhouette based method, the foreground is extracted using background segmentation and then features are extracted from the silhouettes. Bobick & Davis (2001) presented a silhouette based method in which the Motion History Images and Motion Energy Images (MHI, MEI) are used for activity recognition. These MEI and MHI are the images extracted from the video frames and these images are then stacked so as to preserve the temporal content of the activity. Chaaraoui et al. (2014) proposed a method of HAR system that uses the optimized parameters and human silhouette is considered as the basic entity for the feature estimation. The optimized parameters are evaluated using evolutionary computation. Weinland et al. (2007) worked on matching template techniques in which the region of interest (ROI) is divided into a fixed spatial or temporal grid due to which, the effect of noise present in an image and viewpoint

ACCEPTED MANUSCRIPT

variance can be reduced significantly. Thurau & Hlavac (2008) used a histogram of oriented gradients based approach to represent activity, and also concluded that silhouettes are the best information unit for representing human activity. Wu & Shao (2013) proposed modified bags-of- words model called as bag of correlated poses using the advantages of global and local features. They addressed the problem of losing geometric information in bag of visual words representation, which generally is implemented using k-mean clustering algorithm. In these methods, it has been observed that the holistic approach model results in high dimensionality (Agrawal & Ryoo, 2011) of the descriptor; hence there is a need of dimensional reduction techniques for efficient recognition. The PCA is a popular linear dimensionality reduction technique that has been widely used for dimension reduction and classification purpose in activity recognition (Masoud & Papanikolopoulos, 2003; Olivieri, Conde, & Sobrino, 2012). The low dimensional map feature sets are efficiently classified using various linear and nonlinear classifiers. Gorelick et al. (2007) used a nonlinear classification approach for human activity using K-NN along with Euclidian Distance on the global features of the silhouette. Batra et al. (2008) used the nearest neighbour classification approach on local features computed in the form of histograms of code words. Another classification approach that is widely used for human activity recognition based upon local features is SVM and used by Cao et al., 2009; Laptev et al., 2007; and Schuldt et al., 2004. Laptev et al. (2007) used the SVM as well as a KNN classifier to classify human activity and showed that SVM gives better accuracy than KNN, but some factors confine their performances like the interclass similarity and intraclass dissimilarity.

Based upon the analysis of earlier state-of-the art methods on human action and activity recognition, we have captured the problems and listed a layout of the solutions of these problems as follows:

- It is observed that the holistic representation of human activity requires an efficient method for the extraction of silhouette from the video sequence. Usually, foreground segmentation is done using background modelling and background subtraction, but it is not always possible to get good results due to inaccurate modelling of the background. Hence, in this work we have used a texture based segmentation approach to extract the silhouette in context of human activity recognition.

ACCEPTED MANUSCRIPT

- The problem of losing geometrical information in the bags of visual word is addressed by selecting key poses of the human silhouettes. Further, to describe the silhouette information, we have proposed a simple scheme which preserves the spatial change in the human silhouette over time.
- The performance of a classifier reduces when the activities have interclass similarity and intraclass dissimilarity. Therefore, to improve the classification of HAR system, we have constructed a hybrid classification model with the combination of "SVM-NN" classifiers.

3. Proposed Framework

Our approach is based on the silhouette of the human body which is extracted from the video sequence of the activity by segmentation techniques. The segmented silhouette is preprocessed to improve its quality for the feature extraction. Features generated from different silhouette images are then arranged in a representable form. Further, dimension reduction, and classification are used. The workflow diagram of the proposed framework is depicted in Fig. 1 and description of each block is presented in following subsections.

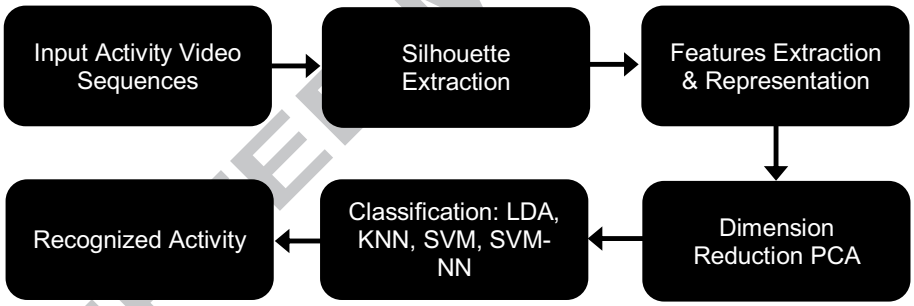


Fig. 1 Workflow diagram of proposed framework

3.1 Extraction of Silhouette using Texture Information

For human activity recognition, the background subtraction is considered as the fundamental challenge of vision based activity recognition. The other challenges which make the task challenging are illumination change, dynamic background, shadows, video noise, etc. (Brutzer, Hoferlin, & Heidemann, 2011). In background subtraction, the fundamental concept is to construct and update the model of the background scene, and foreground object pixels are detected if they differ from the background model up to a certain limit. In the past, Gaussian Mixture Model (GMM) and Local Binary Pattern (LBP) based background models are widely used. In GMM,

ACCEPTED MANUSCRIPT

numerous Gaussian mixture distributions are used to demonstrate each pixel, where each Gaussian distribution characterizes the intensity of pixel distribution over time. For the noisy video sequence, the parameter estimation is unpredictable in case of assumption of Gaussian distribution. Hence, it can be concluded that this assumption is not always true. LBP is a very efficient textural operator, which labels the pixels of the image by thresholding the neighbourhood of each pixel, resulting in a binary pattern. Initially, Haralick et al. (1973), proposed a textural feature based segmentation approach using Gray Level co-occurrence matrix (GLCM). Thereafter, numerous textural feature based segmentation techniques have been proposed (Chua, Wang, & Leman, 2012; Komorkiewicz & Gorgon, 2013; Tahir, Bouridane, & Kurugollu, 2005). A robust textural feature based approach using LBP for background subtraction model is proposed by Chua et al. (2012), in which, they have demonstrated that the textural feature based segmentation method gives more effective results for video surveillance applications. Textural feature based fast and efficient segmentation approach using Gray level co-occurrence matrix (GLCM) has been implemented on FPGA (Komorkiewicz & Gorgon, 2013; Tahir, Bouridane, & Kurugollu, 2005) for different real life applications. The realization of texture based algorithm in real life applications has encouraged us to use textural feature based background subtraction in this work.

A method for describing different textures was originally presented by Haralick et al. (1973). They proposed a matrix called Gray-Level Co-Occurrence Matrix, which allows describing texture based on differences in intensity in different directions and used 14 different features for classification of different textures in the image. Entropy is one the most important parameter that describes the texture information in an image and it can be expressed as:

$$\zeta = \sum_i \sum_j \rho(i,j) \log(\rho(i,j)) \quad (1)$$

Where $\rho(i,j) = \frac{M(i,j)}{\sum_{i,j} M(i,j)}$ is the probability density function; here i and j are indices to the co-occurrence matrix M. The entropy of the image is used to describe its complexity. Higher the entropy, higher is the complexity of the image. An entropy filter is generated in an image to represent the different textures present in the image. The filter matrix is generated for a pixel and its entropy is calculated in 9×9 neighbourhood. Converting this filter matrix into binary form with some thresholding

ACCEPTED MANUSCRIPT

gives an image with white spots at different areas. For a two-textured image, one part contains spots of the same size, and the same is true for the other parts. Removing one such part gives the mask for getting a human blob. Applying this mask over the raw image provides us with the silhouette from this raw image as shown in fig. 2.

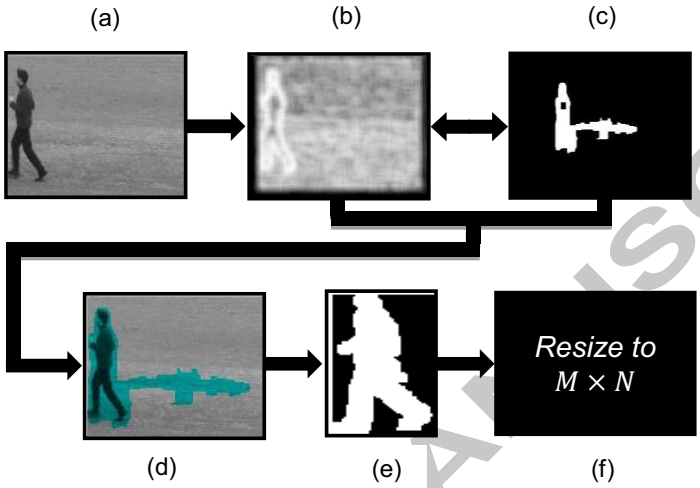


Fig. 2. The flow diagram of silhouette extraction: (a) Original frame of Image, (b) Texture Mask, (c, d) Images selected by rough mask, (e) Silhouette extracted Image from original frame, (f) Final resized image.

The segmented image may contain different white contours, but all of them are not human silhouettes. By comparing the size of these contours one can find the contour with the largest area which is fortunately, the human silhouette. This part of the image is selected and classified as human silhouette. As in Fig. 2, two parts are shown that have the same texture, but human part has the larger area and therefore it has been selected as a silhouette.

3.2 Feature Extraction

Recently, it has been observed that the concept of visual recognition in static images (Guo & Lai, 2014) has been successfully extended to video sequences representation. The various methods (Weinland, Ronfard, & Boyer, 2011), used for representing human activities are: the feature detectors, feature descriptors, bag-of-features representations, and local features based on voting for localization. The feature extraction is the prime step for analysis of video sequences and extracted features must be robust, and invariant against the variation of recording conditions,

ACCEPTED MANUSCRIPT

body pose, etc. All the analysis is done over the feature set collected and based on that, one can find the desired results by applying different techniques. Subsequent sections describe about the key pose selection, and feature extraction and representation.

a) Extracting Key Poses of the Frames

In general some of the video frames do not contain any information content about the object. Consider a person who is 'walking' and the camera is still, he will pass in front of the camera for a short time, and most of the time frames do not have any content in view of 'human blob.' To select the frames, which have maximum information content, the key frames are extracted and are used for the purpose of feature extraction. The Fig.3 depicts the mechanism that is involved in the extraction of key frames out of a large number of frames present in a video.

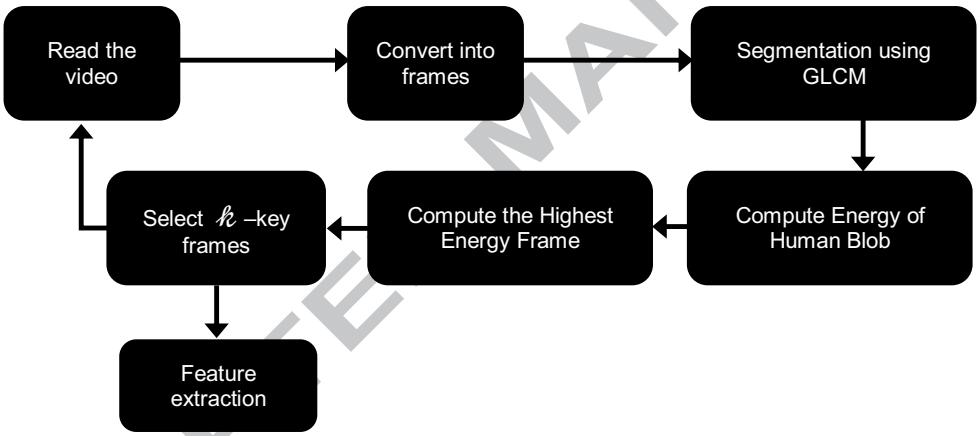


Fig. 3 Flow diagram of selecting key poses frames

The k –key frames, which have significant energy value as compared to the highest energy value of the frame, are chosen for further processing, and energy of the frame is calculated using equation 2. These k –key frames are kept in a timed sequence with reference to the highest energy frame and by this arrangement of key frames the spatial change of the shape with respect to time is maintained. Computation of these key frames is extremely robust in discriminating different human activity due to their ability to discriminate both spatial as well as temporal information.

The timed sequence silhouette frames are divided into a number of cells, and

ACCEPTED MANUSCRIPT

each cell contains different number of white pixels. To maintain the uniformity, the size of the frames has to be fixed. The difference in the size of the obtained silhouette may give different information and may lead to misclassification or error. The silhouettes of person extracted from earlier steps are not of the uniform size and therefore, it is necessary to resize the images.

b) Cells Formation

The resized image of $M \times N$ contains total pixels equivalent to $\mathbb{N}$. It is divided into a grid of $\mathcal{U} \times \mathcal{V}$ images as depicted in Fig. 4. Since the image is converted into binary form, therefore the white pixel can be calculated in the cell and the number of white pixel is used as a feature for this particular cell or grid.

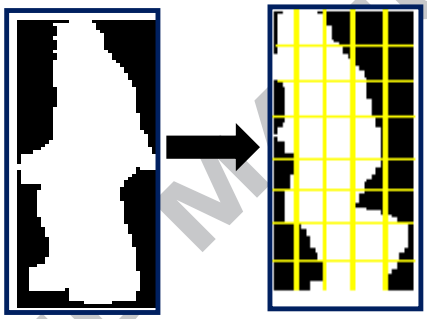


Fig .4 Formation of Cells using key frames

c) Feature Extraction and Representation

Consider a segmented video of an activity that contains a finite number $(\mathfrak{N})$ of frames, represented as $I_t(x, y)$, where t represents the frame number i.e. $t \in \{1,2,3, \ldots \ldots, \mathfrak{N}\}$ and x, y are the dimensions of the frames. To maintain uniformity, the next step is to resize each frame to a size of $M \times N$.

For effective and efficient representation of activity, only key poses of the frames are chosen and these are the frames that have higher energy in the video sequence compared to other frames. The energy of a frame is calculated as:

$$U_t = \sum_{i=1}^{M} \sum_{j=1}^{N} \|I_t(i,j)\|^2 \quad (2)$$

For the selection of key poses of the silhouette frames, a sequential search operation is applied up to a certain number of frames to find highest energy value of silhouette frame among all frames $(U_1, U_2, U_3, U_4, \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots U_{\mathfrak{N}})$. The

ACCEPTED MANUSCRIPT

highest energy value frame is considered as the reference frame and the frames which have significant energy value as compared to the highest energy value frame are selected as k-key frames of the silhouette. Now, each key frame is divided into cell images $C_i(x,y)$, where each cell is of size $u \times v$, and the total number of cells therefore are calculated as:

$$\frac{M}{u} \times \frac{N}{v} = N_c \tag{3}$$

Where N_c is total number of cells in the key frames and denoted as ($C_1, C_2, C_3, C_4, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, C_{N_c}$) .Since silhouette images are binary images, they may contain only white or black pixels and the number of white pixels are counted as:

$$w_i = count\{C_i(x,y)\},\ where\ i = 0,1,2, \ldots\ldots\ldots\ldots N_c. \tag{4}$$

Where w_i represents the number of white pixels in the i^{th} cell, the counted number of pixels in a cell of one frame are arranged in such a manner that retains the time sequence of the action and can be represented as:

$$f_i = \{w_1, w_2, w_3, w_4, \ldots\ldots\ldots\ldots w_{N_c}\}, where\ i = 1,2,3, \ldots\ldots\ldots k. \tag{5}$$

Where f_i contains the white pixel count of the i^{th} key frame. Thus, the feature vector of an activity video sequence is expressed as:

$$F_i = [f_1, f_2, f_3, f_4, \ldots\ldots\ldots\ldots\ldots f_k], where\ i = 1,2,3,4, \ldots\ldots\ldots\ldots\ldots\ldots V_T. \tag{6}$$

where V_T is the total number of videos in the data set. Substituting the equation 5 in equation 6, then feature vector for a single video of activity is:

$$F_i = \left[\underbrace{w_1, w_2, w_3, w_4, \ldots\ldots w_{N_c}}_{First\ Frame}, w_1, w_2, w_3, w_4, \ldots\ldots\ldots \underbrace{w_1, w_2, w_3, w_4, \ldots\ldots w_{N_c}}_{Last\ Frame\ i.e.k^{th}}\right] \tag{7}$$

Similarly, the feature set of a dataset which contains V_T number of videos of all classes can be represented as:

ACCEPTED MANUSCRIPT

$$F_v = \begin{pmatrix} F_1 = & \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{First\ Frame}, w_1, w_2, w_3, w_4, \dots \dots \dots, \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{Last\ Frame\ i.e.\ k^{th}} \\ F_2 = & \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{First\ Frame}, w_1, w_2, w_3, w_4, \dots \dots \dots, \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{Last\ Frame\ i.e.\ k^{th}} \\ & \vdots \\ F_{V_T} = & \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{First\ Frame}, w_1, w_2, w_3, w_4, \dots \dots \dots, \underbrace{w_1, w_2, w_3, w_4, \dots \dots w_{N_c}}_{Last\ Frame\ i.e.\ k^{th}} \end{pmatrix} \quad (8)$$

where v corresponds to the videos of all classes. One of the important aspects of the feature set is its dimension. If the dimension is high, then it must be reduced to a lower dimension for efficient and speedy classification. The dimension of the feature set F_v can be determined by determining the dimension of feature vector representing a single video in equation 7. It is determined to be $N_f = N_c \times k$ and after concatenation the dimension is written as $1 \times N_f$. Hence, the dimension of the final feature set F_v is determined as $V_T \times N_f$. Finally, in order to recognize actions, these feature vectors and their labels are passed to the classifiers.

3.3 *Classification Model*

Usually, the computed data of feature set have the correlated and uncorrelated data together and the correlated feature set of different classes makes the classification complex and slow. Hence, for efficient and fast classification, the feature set must have uncorrelated data since the uncorrelated feature set has lower dimension. Therefore, for easy handling of data and improved classification, the dimension of feature set must be reduced. The PCA (Jolliffe, 1986) is a popular method for the reduction of dimension of the feature set by maximizing the variance of the feature set and mapping the feature set into a lower dimensional space.

The significant lower dimension data is classified by training and testing of the feature set. There are two types of classifiers used for the human activity recognition, which are i) Linear and ii) Nonlinear. We have used one linear and two nonlinear classifiers. The linear classifier is the LDA, and nonlinear classifiers are: KNN, and SVM. In addition to these one more nonlinear classification model is proposed using the combination of SVM and KNN.

ACCEPTED MANUSCRIPT

Originally, the concept of Linear Discriminant Analysis (LDA) was proposed by Fisher (1936), also known as Fisher's Discriminant Analysis. It provides an improved classification of various classes of data by maximizing the margin between dissimilar classes and minimizing the margin within the same classes.

The KNN classifier (Cover & Hart, 1967) selects the K-closest samples of training feature set to the new instance, and the nearest class having highest voting is mapped to test instance. One of the biggest advantages of this classifier is its non-parametric nature. It does not require any assumptions and easily classifies the data even in higher dimension space.

Support vector machine (SVM) is a machine learning classifier based on the structural risk minimization principle (Vapnik, 1999). It is one of the widely used classifiers for classification of human activity (Hsu & Lin, 2002). The term support vectors are features of the samples, which are close to the hyperplane of an SVM. Therefore, the determination of the location of most important data is near the hyperplane and it is formed by the set of lines drawn between the class samples.

To enhance the recognition accuracy, a hybrid form of "SVM-NN" classification model is introduced which utilizes the individual benefits of SVM and NN model. The proposed flow of layout of the classification model of "SVM-NN" is as depicted in Fig. 5.

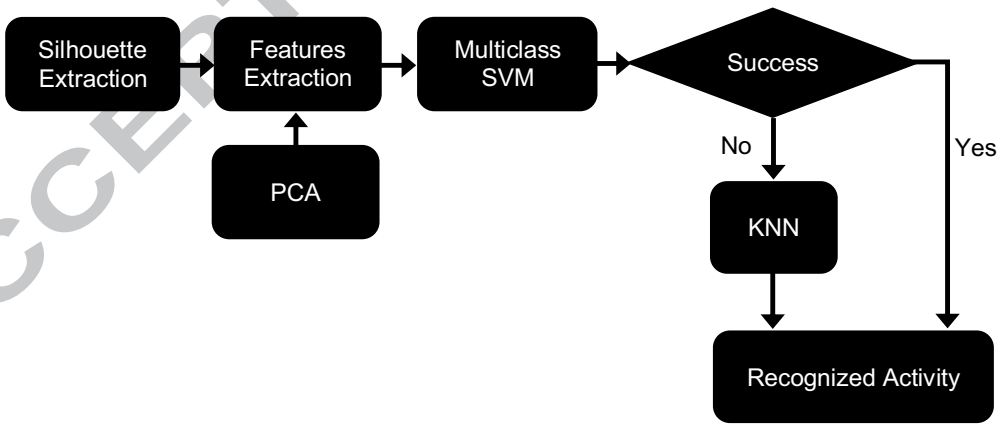


Fig. 5 Showing Hybrid Classification Model of "SVM-NN"

In this model, initially the SVM is used to classify the input feature set and in these feature sets, some are correctly classified and some are wrongly classified.

ACCEPTED MANUSCRIPT

The wrongly classified feature sets lie near the separating hyperplane and are the support vectors. Further, these support vectors are classified using NN, and are considered as representative points.

4. Experimental Results and Discussion

In order to assess the effectiveness of the proposed approach, we have conducted experiments on three public benchmarks, the Weizmann data set (Gorelick, Blank, Shechtman, Irani, & Basri, 2007), the KTH data set (Schuldt, Laptev, & Caputo, 2004), and the Ballet data set (Fathi & Mori, 2008). The assessments include a range of variations resulting from different resolutions, lighting conditions, occlusions, background disorder, and uneven motions. In this experiment, for the representation of all videos of the data sets, 31 key frames of size 40×25 are used to represent an activity sequence and each frame has 40 number of cells, which is obtained by considering the size of each cell 5×5. Hence, the dimension of a silhouette is $31 \times 40 = 1240$, and after concatenation the silhouette feature vector is represented as 1×1240. To evaluate the outcome of action classifications, a leave-one-out (LOO) cross validation scheme is opted for all the data sets. For each data set, we have used three machine learned classifier (LDA, KNN, SVM). In addition to these, to enhance the recognition accuracy, a hybrid "SVM-NN" classification model is proposed. The average recognition accuracy (ARA) is computed using equation 9.

$$\mathrm{ARA} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \times 100 \text{ (In percentage)} \quad (9)$$

where TP, TN, FP and FN are the number of true positive, true negative, false positive, and false negative, respectively. The highest obtained recognition accuracy of these classifiers are compared with the similar state-of-the-art methods.

Weizmann data set: This dataset was introduced by Gorelick et al. (2007), which consists of 90 videos with a frame rate of 25fps and each frame has a size of 144× 180. In the video sequence, 9 people are performing 10 different actions, categorized as walk, run, jump-jack, bend, jumping forward on one leg, jumping on two legs in the forward direction, jumping in place, sideways jump, one hand wave, two hand wave. This is one of the established benchmarks of the data sets for the human action recognition and most of the earlier methods' (Gorelick, Blank, 2007; Chaaraoui, Pérez & Revuelta, 2013; Wu & Shao, 2013; Goudelis, Karpouzis, &

ACCEPTED MANUSCRIPT

Kollia, 2013; Melfi, Kondra & Petrosino, 2013; Touati & Mignotte, 2014) recognition accuracy is computed on this data set. The sample frames of the data set are shown in Fig. 6.

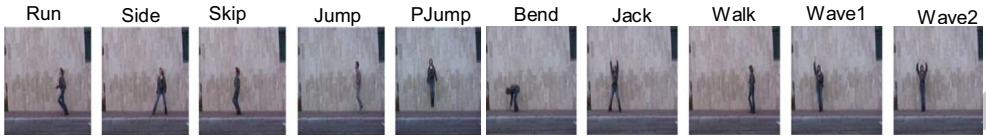


Fig. 6. Sample frames of Weizmann Human Action dataset.

KTH data set: This dataset was introduced by Schuldt et al. (2004) and is a more challenging dataset as compared to the Weizmann dataset. The dataset consists of six basic activities, namely; "Hand-Clapping," "Hand-Waving," "Jogging," "Jumping," "Running," and "Walking". Each activity has 100 videos for four different scenarios in different light conditions, indoor and outdoor conditions. All these video sequences are recorded in a uniform background with a static camera of frame rate 25fps and further down-sampled to the spatial resolution of 160x120 pixels. The recording conditions of the videos in the KTH data set are not stationary and there is a significant amount of camera movement and lighting effects in some cases. Therefore, the silhouette extraction is not straight forward and simple background subtraction method may not suitable. Hence, for the silhouette extraction we have incorporated the texture based segmentation method. The sample images of the data set are depicted in Fig. 7.

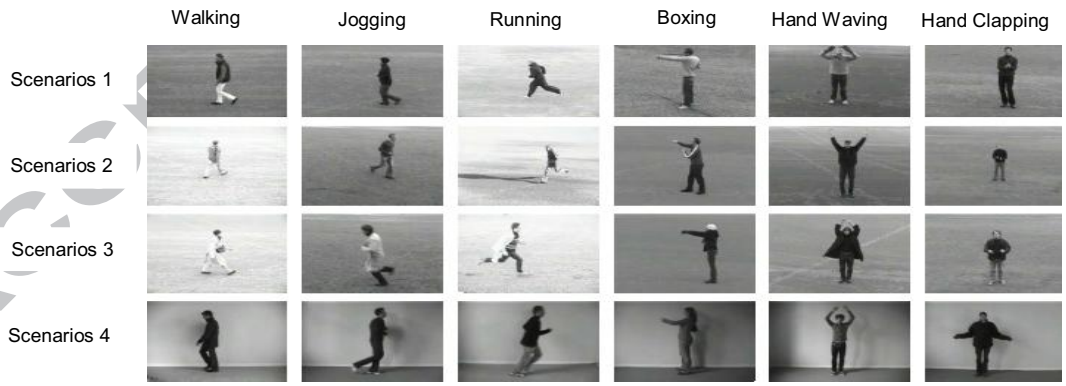


Fig. 7. Sample frames of KTH dataset

Ballet data set: The Ballet Movement action data set (Fathi & Mori, 2008) is one of the complex human action data sets. This data set consist of eight ballet movements performed by three actors and these movements are named as " Hopping (HP),"

ACCEPTED MANUSCRIPT

"Jump (JP)," "Left-to-Right Hand Opening (LRHO)," "Leg Swinging (LS)," "Right-to-Left Hand Opening (RLHO)," "Standing with Hand Opening (SHO)," "Stand Still (SS)" and "Turning (TR),". The data set is highly challenging due to the considerable amount of intra-class dissimilarities in terms of spatial and temporal scale, speed, and clothing. The sample frames of the data set are depicted in Fig. 8.

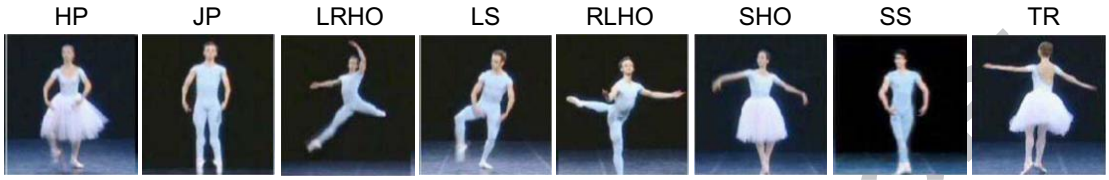


Fig. 8. Images of the Ballet data set depicting eight movement of actions

4.1 Classification Results

The classification results of the proposed approach are depicted in Table 1 on three different data sets using four different classification models, including the proposed one i.e. "SVM-NN". The aim of this depiction is to show the effectiveness of the proposed descriptor as well as the performance of the proposed classification model in comparison with the existing models.

The classification strategy opted for the Weizmann data set is as per Gorelick et al., 2007. From Table 1, the recognition accuracy obtained by LDA is less as compared to the other classifiers. This is due to the similarity between the activities of the data set like running, jumping and walking and therefore activities are very difficult to separate using a linear model of classification. The highest ARA achieved in our experiment is through hybrid "SVM-NN" classification model which is 100%. The hybrid classifier is the combination of two nonlinear classifiers and these classifiers are more adept to inter class similarities and intra class dissimilarities, hence giving the improved result.

For the KTH data set, the variation in recording condition is more as compared to the Weizmann data set due to which the extraction of silhouette in this data set is a difficult task. Simple frame differencing methods for extraction of silhouette may give good results in Weizmann data set, but in the case of KTH, it is very difficult to extract silhouette using these methods. Due to the variation in recording conditions the texture of object is least affected. Hence, we utilize the texture based foreground extraction. The highest ARA achieved on the KTH data set in our experiment is 96.4

ACCEPTED MANUSCRIPT

% as shown in Table 1. From Table 1, it can also been seen that the performance of various classifiers on the KTH data set is increasing.

Table 1

Classification Results on the Data Sets in terms of ARA (%)

Data Sets\ Classifiers	Weizmann	KTH	Ballet	mRA (%)
LDA	*94.4*	*90*	*75*	*86.4*
KNN	*96.6*	*91.7*	*90.2*	*92.8*
SVM	*97.7*	*92.4*	*92.75*	*94.28*
SVM-NN	***100***	***96.4***	***94***	***96.8***

For the Ballet data set, the highest ARA achieved is 94% through SVM-NN classifier as shown in Table 1. The performance of our approach is least on this data set as compared to the Weizmann and KTH data sets, because of the complex motion patterns, which differentiate in execution of the motion from actor to actor. The misclassification error is instigated by the "hopping" as it is confused with a much related activity "jump".

4.2 Comparison of Recognition Accuracy:

Comparison of recognition accuracy is carried out in two stages: first, the performance of proposed "SVM-NN" classifier is compared with others (LDA, KNN, and SVM) and the second stage, the highest average recognition accuracy obtained on the data sets, is compared with the methods of others.

The classifier performance is compared through the mean classification error (MCE), which is computed using Mean Recognition Accuracy (mRA). The MCE of all the classifiers is depicted in the Fig.9, and lowest classification error is found in the case of SVM-NN classifier.

ACCEPTED MANUSCRIPT

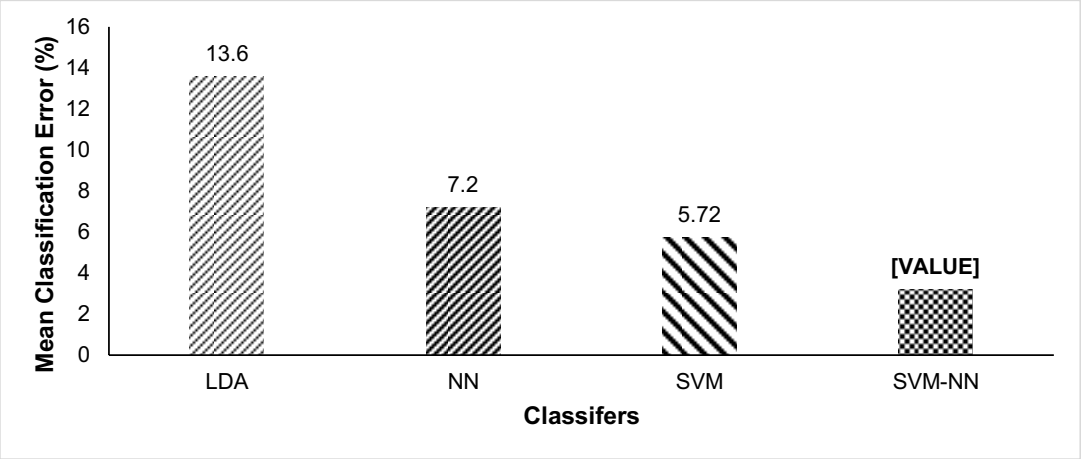


Fig. 9. Showing the comparative Performance of the Classifiers

LDA is a linear classifier, the speed of processing of which is faster than a kernel based classifier, but the speed comes at the cost of efficiency. The MCE is 13.6 %, which is higher than others and hence it can be concluded that it may not be a good classifier for the dataset which has more interclass similarity. The MCE of a nearest neighbour classifier is almost 7.2 %, which is slightly lower than that of LDA and the reason for the lower MCE is the nonlinear nature of the classifier. The SVM is a kernel based trick, and gives a unique solution because the optimality problem is convex. The MCE is reported as 5.7 %, which is lower than the LDA and KNN both. The individual performance of SVM and KNN approach for classification of human activity is already proven (Gorelick et a., 2007; Chaaraoui et al., 2013; Wu & Shao, 2013; Goudelis et al., 2013; Melfi et al., 2013; Touati & Mignotte, 2014; Sadek et al., 2011; Saghafi & Rajan, 2012; Goudelis et al., 2013; Melfi et al., 2013; Rahman et al. 2014; Conde & Olivieri, 2015). Therefore, a hybrid "SVM-NN" classifier is proposed and the performance is measured in terms of MCE (3.2%), which is lowest among all the classifiers used.

The effectiveness of our proposed description methodology of human activity is analyzed by comparison of the highest ARA achieved on each data set with respect to the other methodologies as presented in Table 2, 3, 4. The highest ARA is achieved by the hybrid "SVM-NN" classification model on all the data sets used.

ACCEPTED MANUSCRIPT

Table 2

Comparison of our results with other similar state-of-the-art methods on Weizmann data set

Method	Input	Classifiers	Test Scheme	ARA (%)
Gorelick et al. (2007)	*Silhouettes*	*KNN*	*LOO*	*97.5*
Chaaraoui et al. (2013)	*Silhouettes*	*KNN*	*LOSO*	*92.8*
Wu, & Shao (2013)	*Silhouettes*	*SVM*	*LOSO*	*97.78*
Goudelis et al. (2013)	*Silhouettes*	*SVM*	*LOPO*	*95.42*
Melfi et al. (2013)	*Silhouettes*	*SVM*	*LOO*	*99.02*
Touati & Mignotte (2014)	*Silhouettes*	*KNN*	*LOO*	*92.3*
Proposed Method	*Silhouettes*	*SVM-NN*	*LOO*	*100*

Table 3

Comparison of our results with other similar state-of-the-art methods on KTH data set

Methods	Input	Classifiers	Test scheme	ARA (%)
Sadek et al.(2011)	*Silhouettes*	*SVM*	-	93.30
Saghafi & Rajan (2012)	*Silhouettes*	*KNN*	*LOO*	92.6
Goudelis et al. (2013)	*Silhouettes*	*SVM*	*LOPO*	93.14
Melfi et al. (2013)	*Silhouettes*	*SVM*	*LOO*	95.25
Rahman et al. (2014)	*Silhouettes*	*KNN*	*LOO*	94.49
Conde & Olivieri (2015)	*Images*	*KNN*	-	91.3
Proposed Method	*Silhouettes*	*SVM-NN*	*LOO*	96.70

Table 4

Comparison of our results with other similar state-of-the-art methods on Ballet data set

Methods	Fathi & Mori, (2008)	Wang & Mori (2009)	Guha, & Ward (2012)	Ming et al. (2013)	Iosifidis et al. (2014)	Proposed Method
ARA (%)	*51*	*91.3*	*91.1*	*90.8*	*91.1*	*94*

Tables 2 & 3, show the comparison of our result achieved through “SVM-NN” classification model with the similar state-of-the-art methods on Weizmann and KTH data set, respectively. The test methodologies used in these methods are Leave-

ACCEPTED MANUSCRIPT

One-Out (LOO), Leave-One-Person-Out (LOPO), and Leave-One-Sequence-Out (LOSO), which are fairly similar to each other. Hence, the comparison on these two data sets is fair, because the experimental setup used in these techniques is similar to that in the proposed one. As it is seen in Table 2, the ARA of 100% is achieved on Weizmann data set, which is higher than the other methods that use the SVM and KNN classification models. The reason for this high accuracy is the quality of silhouette extraction and effective representation, and capability of classifier to deal with intraclass variation among the activities. Similarly, as it is seen in table 3, the ARA is 96.7%, which is again higher than the other methods that use the similar classification model.

Table 4, gives the comparison of our result with that of five earlier works which use the Ballet data set. We have used the similar experimental setup as used in Fathi & Mori (2008), Wang & Mori (2009), Ming et al. (2013), & Iosifidis et al. (2014), thus this comparison is fair. The work of Guha & Ward (2012) uses different experimental setup as compared to ours. Hence, this comparison may not be a right one, but for the given complexity of the database, the higher ARA is obtained successfully by our technique, which is heartening.

From these experimental results, few interesting observations can be made:

- Our proposed approach of activity representation and classification model significantly outperforms many of the existing silhouette based activity recognition methods as can be seen from Table 2, 3, and 4.
- Our classification model performs better than the LDA, KNN, and SVM as can be seen by Fig. 9, in which the least error is achieved through SVM-NN.
- It can be seen that improvements achieved for the KTH and Ballet data set are more significant because these two data sets have challenging environmental conditions and significant intraclass variations in terms of speed, spatiotemporal scaling, zooming in, zooming out, clothing, etc. and these are directly related to the input data. The silhouette extraction on Weizmann data set is comparatively easy and accurate as compared to the KTH and Ballet data set due to less variation in recording conditions.
- As the number of key poses increases, the complexity also increases, and it does not give a significant increase in recognition accuracy. On the other hand increasing the number of cells, gives marginal increase in the

ACCEPTED MANUSCRIPT

effectiveness but results in a higher dimension.

5. Conclusion

In this paper, a vision based human activity recognition system exploiting the key poses of the human silhouettes is presented. The problem of less recognition rate under the variant environmental conditions has been addressed by employing: 1) Accurate human silhouette extraction through texture based background subtraction approach 2) Simple and effective representation of human silhouettes by means of grids and cells. 3) An efficient hybrid classification model of "SVM-NN". The effectiveness of the proposed approach is tested on three public data sets through LDA, KNN, SVM, and "SVM-NN" classification models and ARA of these models are measured. The success of these classification models is assessed using MCE and it is observed that the hybrid classification model i.e. the "SVM-NN" gives the least classification error. The overall performance of the proposed approach is found to be comparatively more effective. The parameters used for feature representation are simple and the computation is easy. The three data sets used here vary in terms of lighting conditions, indoor and outdoor environment, zoom in, zoom out, and hence it can be concluded that the proposed approach is robust under such conditions.

Despite the satisfactory results, some concerns have cropped up: (i) It is imperative that only one person is in the video sequence, (ii) Some parameters like the number of key poses, size of grids and cells can be further optimized (iii) This approach is less effective, when object is occluded.

For the future, one can optimize these parameters further so that more effective and accurate representation is conceivable. The same approach may be extended for other avenues of research like Human Style Recognition, Hand Gesture Recognition, Facial Recognition etc.

An expert and intelligent system can be developed using this approach for variety of applications such as: (i) Telemedicine system for providing assistance to Parkinson disease patients (ii) An intelligent system, which can monitor the elderly person for abnormal activity, and (iii) Intelligent surveillance system, which can raise an alarm during theft, robbery etc. (iv) A system which can coach athletes to improve their techniques by providing correct assistance e.g. golf swing, cricket swing etc.

References

ACCEPTED MANUSCRIPT

Agrawal, J. K., & Ryoo, M. S. (2011). Human Activity Analysis: A Review. *ACM Computing Survey* 43 (3), 16-43.

Batra, D., Chen, T., & Sukthankar, R. (2008). Space–time shapelets for action recognition. *Proc. of the Workshop on Motion and Video Computing,* 1–6.

Bobick, Aaron F., & Davis, James W. (2001). The recognition of human movement using temporal templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence,* 23 (3), 257–267.

Brutzer, S., Hoferlin, B., Heidemann, G. (2011). Evaluation of Background Subtraction Techniques for Video Surveillance. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Cao, D., Masoud, O.T., Boley, D., & Papanikolopoulos, N. (2009). Human motion recognition using support vector machines. *Computer Vision and Image Understanding*, 113, 1064–1075.

Chaaraoui, A. A., Pérez, P.C., & Revuelta, F. F. (2012). A review on vision techniques applied to Human Behaviour Analysis for Ambient-Assisted Living. *Expert systems with applications, 39, 10873-10888.*

Chaaraoui, A.A. & Revuelta, F. F. (2014).Optimizing human action recognition based on a cooperative coevolutionary algorithm. *Engineering Applications of Artificial Intelligence, 31, 116-125.*

Chaaraoui, A.A., Pérez, P. C. & Revuelta, F.F. (2013).Silhouette-based human action recognition using sequences of key poses. *Pattern Recognition Letters, 34, 1799-1807.*

Chakraborty, B.; Holte, M.B.; Moeslund, T.B.; & Gonzàlez, J. (2012). Selective spatio-temporal interest points. *Computer Vision and Image Understanding,* 116(3), 396-410.

Cheema, S., Eweiwi, A., Thurau, C., Bauckhage, C., (2011). Action recognition by learning discriminative key poses. *Proc. of IEEE International Conference on Computer Vision Workshops (ICCV Workshops),* 1302–1309.

Conde, I.G., & Olivieri, D. N. (2015). A KPCA spatio-temporal differential geometric trajectory cloud classifier for recognizing human actions in a CBVR system. *Expert Systems with Applications, http://dx.doi.org/10.1016/ j.eswa.2015.03.010.*

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13, 21–27.

Chua, T.W., Wang, Y., & Leman, K. (2012). Adaptive Texture-Color based Background Subtraction for Video Surveillance. *Proc. of 19TH IEEE International conference on image processing (ICIP).*

Dollar, P., Rabaud, V.,Cottrell, G.,& Belongie, S. (2005). Behavior Recognition via Sparse Spatio-Temporal Features. *Proc. of 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance,* 65-72.

Everts, I., Gemert, J.C.V., & Gevers, T. (2014). Evaluation of Color Spatio-Temporal Interest Points for Human Action Recognition. *IEEE Transactions on Image Processing,* 23(4), 1569-1580.

Eweiwi, A., Cheema, S., Thurau, C., Bauckhage, C. (2011). Temporal Key Poses for

ACCEPTED MANUSCRIPT

Human Action Recognition. *Proc. of IEEE International Conference on Computer Vision, 1310-1317.*

Fathi, A. & Mori, G. (2008). Action Recognition by Learning Mid-level Motion Features. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition.*

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. Annals of Eugenics, Vol. 7, pp. 179-188.

Gorelick, L., Blank, M., Shechtman, E., Irani, M., & Basri, R. (2007). Actions as space–time shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29 (12), 2247–2253.

Guha, T. & Ward, R.K. (2012). Learning Sparse Representations for Human Action Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(8), 1576-1588.

Guo, G., & Lai, A. (2014). A survey on still image based human action recognition. *Pattern Recognition,* 47, 3343–3361.

Goudelis, G., Karpouzis, K.,& Kollia, S.(2013).Exploring trace transform for robust human action recognition. *Pattern Recognition*, 46, 32-38-3248.

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Trans. on Systems, Man, and Cybernetics*, 6, 610-621.

Hsu, C. W., & Lin, C. J. (2002). A Comparison of Methods for Multiclass Support Vector Machines. *IEEE Trans. on Neural Network*, 13(2), 415-425.

Iosifidis, A., Tefas, A., Pitas, I. (2014). Discriminant Bag of Words based representation for human action recognition. *Pattern Recognition Letters*, 49, 185-192.

Jargalsaikhan, I., Little, S., Direkoglu, C. & O'Connor, N.E. (2013). Action recognition based on sparse motion trajectories. *Proc. of IEEE International Conference on Image Processing,* 3982-3985.

Jolliffe, I. (1986). Principal Component Analysis. New York: Springer-Verlag.

Komorkiewicz, M. & Gorgon, M. (2013). Foreground Object Features Extraction with GLCM Texture Descriptor in FPGA. *Proc. of IEEE Conference on Design and Architectures for Signal and Image Processing (DASIP).*

Laptev, I., Caputo, B., Schuldt, C., & Lindeberg, T. (2007). Local velocity-adapted motion events for spatio-temporal recognition. *Computer Vision and Image Understanding,* 108 (3), 207–229.

Laptev, I. (2005). On Space-Time Interest Points. International Journal of Computer Vision 64(2/3), 107–123.

Masoud, O., & Papanikolopoulos, N. (2003). A method for human action recognition. *Image and Vision Computing*, 21 (8), 729–743.

Melfi, R., Kondra, S., & Petrosino, A. (2013). Human activity modeling by spatio temporal textural appearance. *Pattern Recognition Letter*, 34, 1990-1994.

Ming, X.L., Xia, H.J.,Zheng, T.L.(2013). Human action recognition based on chaotic invariants. *Jounral of South Central University,* 20, 3171-3179.

Olivieri, D. N., Conde, I. G., & Sobrino, X. A. V. (2012). Eigen space based fall detection, and activity recognition from motion templates and machine learning.

ACCEPTED MANUSCRIPT

Expert Systems with Applications, 39, 5935–5945.

Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 28, 976–990.

Rahman, S. A., Song, I., Leung, M.K.H., Lee, I., & Lee, K. (2014). Fast action recognition using negative space features. *Expert Systems with Applications*, 41, 574–587.

Ryoo, M.S. & Aggarwal, J. K., (2009). Spatio-temporal relationship match: Video structure comparison for recognition of complex human activities. *Proc. of the IEEE International Conference on Computer Vision*, Los Alamitos, CA.

Sadek, S., Hamadi, A. A., Elmezain, M., Michaelis, B., & Sayed, U. (2012). Human Action Recognition via Affine Moment Invariants. *21st International Conf. on Pattern Recognition,* 218 – 221.

Saghafi, B., & Rajan, D. (2012). Human action recognition using Pose-based discriminant embedding. *Signal Processing: Image Communication*, 27, 96-111.

Schuldt, C., Laptev, I., & Caputo, B. (2004). Recognizing human actions: a local SVM approach. *Proc. of the International Conf. on Pattern Recognition*, 3, 32–36.

Tahir, M.A., Bouridane, A., & Kurugollu, F. (2005). An FPGA Based Coprocessor for GLCM and Haralick Texture Features and their Application in Prostate Cancer Classification. *Analog Integrated Circuits and Signal Processing,* 43, 205–215.

Thurau, C., & Hlavac, V. (2008). Pose primitive based human action recognition in videos or still images. *Proc. of the Conf. on Computer Vision and Pattern Recognition*, 1–6.

Touati, R. & Mignotte, M. (2014). MDS-Based Multi-Axial Dimensionality Reduction Model for Human Action Recognition. *Proc. of IEEE Canadian Conference on Computer and Robot Vision*, 262-267.

Vapnik, V. N. (1999). An Overview of Statistical Learning Theory, *IEEE Trans. On Neural Network*, 10(5), 989-999.

Vishwakarma, S., & Agrawal, A. (2012). A survey on activity recognition and behavior understanding in video surveillance. *Visual Computer*, 29, 983-1009.

Wang, Y., & Mori, G.(2009).Human Action Recognition by Semilatent Topic Models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 31(10), 1762-1774.

Weinland, D., Ronfard, & R., Boyer, E. (2011). A survey of vision-based methods for action representation, segmentation, and recognition. *Computer Vision and Image Understanding*, 115, 224–241.

Weinland, D., Boyer, E., Ronfard, R. (2007). Action Recognition from Arbitrary Views using 3D Exemplars, *IEEE International Conf. on Computer Vison*, 1-7.

Wu,D., & Shao,L. (2013).Silhouette Analysis-Based Action Recognition Via Exploiting Human Poses. *IEEE Transactions on Circuits and Systems for Video Technology,* 23(2), 236-243.

Ziaeefar, M., & Bergevin R. (2015). Semantic human activity recognition: A literature review. *Pattern Recognition, http://dx.doi.org/10.1016/j.patcog.2015.03.006.*

ACCEPTED MANUSCRIPT

Highlights

- A new approach for the recognition of human activity using silhouette is proposed.
- The effectiveness of the proposed approach is measured using various classifiers.
- A new hybrid classification model is proposed to boost the recognition accuracy.
- The minimum classification error is achieved through a hybrid classification model.

ACCEPTED MANUSCRIPT

Modified Hysteresis Switching Scheme for 3-Phase Voltage Source Inverter

Ankit Yadav
Department of Electrical Engineering
Institute of Engineering &Technology (JKLU)
Jaipur, India
ankit.jnvf@gmail.com

Dr. M. M. Tripathi
Department of Electrical Engineering
Delhi Technological University
New Delhi, India
mmtripathi@dce.ac.in

Abstract –This paper presents a new switching scheme for 3-phase Voltage source inverter to tackle the harmonics problem in off-grid power system. The proposed strategy is modification over the existing strategy of hysteresis band current control method. The proposed switching strategy for 3-phase inverter system is an upgrade of the scheme that has been tested already with 1- phase system. Simulation with linear as well as nonlinear loads has been performed using new scheme as well as PWM scheme. The results obtained have been compared with results of PWM scheme and the proposed scheme has shown much improved outcomes compared to the PWM scheme. The comparatively higher output voltage and current also ascertain the usefulness of new proposed switching methodology.

Keywords – DC-AC Inverter, Feedback, Gate pulses, PWM (pulse width modulation), Switching

I. INTRODUCTION

In the current era of energy crisis, the renewable energy sources have emerged as necessity for sustainable energy. The renewable energy can be utilized through off grid as well as grid connected mode of operations. In both modes of operation the inverter plays a significant role as an intermediate device between the renewable energy source and the load. The quality of power supplied is dependent on the inverter operation. Along with converting DC to AC, inverters also injects the harmonics in the power system. These harmonics distort the power supply, and may cause serious harmful effect on any sensitive load, if connected. The quality of power and harmonic injected by inverter is function of the switching strategy being used for operating the inverter device. Standards like as IEEE 1547, IEEE 519 have been laid down from IEEE and other institutions to control the power quality and harmonic content in the supply. Several attempts have been made by researchers worldwide [1]-[8] to tackle the harmonics problem. Some schemes like PWM, delta modulation, hysteresis control etc. have shown significant results and are being used nowadays in many industrial applications[9]-[11].

This paper presents a new switching strategy for 3- phase voltage source DC–AC inverter. The new scheme is modification over the existing hysteresis band current control. The new scheme has been tested already with 1- phase H–bridge voltage source DC–AC inverter [12]. In this paper a new modified scheme is presented for three phase system. The new strategy has been tested with linear and nonlinear loading conditions and analyzed with respect to PWM scheme under the similar loading conditions. The newly developed scheme has shown significant improvement over PWM scheme.

The section-II is dedicated to modified switching scheme for 3-phase. Section-III presents the implementation of the developed scheme. Section-IV is simulation results. Conclusion belongs to section-V and at end section-VI contains the references.

II. SWITCHING SCHEME FOR 3-PHASE INVERTER

The proposed switching scheme has been developed for 3-phase voltage source inverter. The circuit diagram of 3-phase inverter is shown in Fig. 1. It contains three legs, each with two switches. All the development and testing has been performed using the same circuit diagram. The IGBT is selected as power electronics switch. The switches S1 to S6 are provided with GATE signal through G1 to G6 respectively. These pulses follow the applied switching scheme and control the inverter accordingly. The new scheme is discussed below in detail.

A. New Switching Strategy

The block diagram of the new switching scheme for three-phase inverter as shown in fig.2 is a modification of same scheme used for single-phase system. The value of load currents are measured through ammeter and then their values are passed to switching circuit as feedback. Thus, the feedback signals are images of original load current waveforms.

The reference signals for this scheme are sinusoidal signal as defined by eq. (1).

$$R_a = R_{a_{max}} \sin(wt) \quad (1.a)$$

$$Y_a = Y_{a_{max}} \sin(wt - wT_m/3) \quad (1.b)$$

$$B_a = B_{a_{max}} \sin(wt - wT_m/3) \quad (1.c)$$

Where R_a, Y_a, B_a are reference signals for R, Y, B phases respectively. Ra_{max}, Ya_{max}, and Ba_{max} are the maximum value of reference signals respectively.

All the three reference signals are images of the desired load current in the respective phases. The feedback signal also can be named as R_b, Y_b and B_b respectively. Generating switching pulses for the three-phase inverter follows the same concept as single-phase inverter.

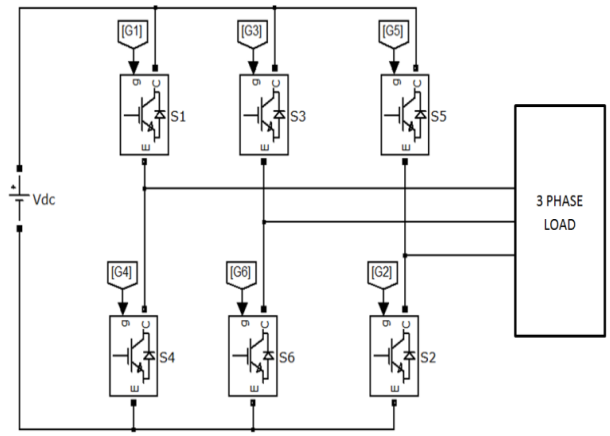


Fig.1:– circuit diagram of 3 – phase voltage source inverter.

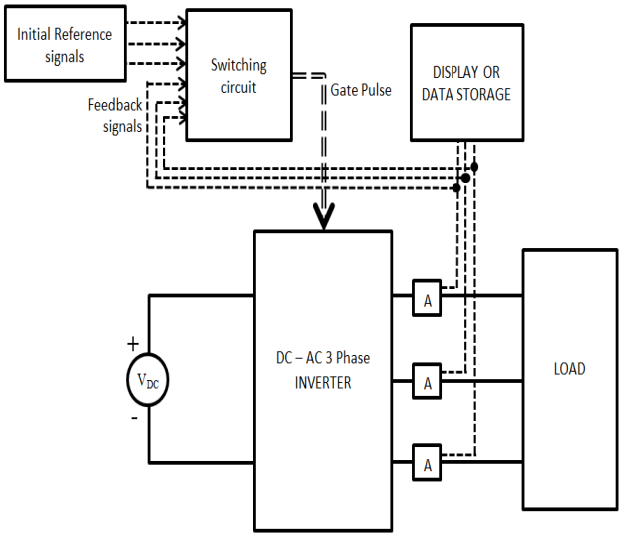


Fig.2:- Block diagram for modified switching scheme.

B. Reference signal

The uniqueness of the new strategy lies in its reference signals. Instead of having a fixed reference signal and scaling the modulating signal to that range the, a new concept of variable reference signal has been introduced. This reduces the calculation size and eliminates the circuitry to scale the modulating signal. Each reference signal includes two parts.

1) Initial reference signals
2) Updated reference signal

III. IMPLEMENTATION

S–Function block of SIMULINK facility of MALAB R2011b has been used for developing the switching logic. The program for switching logic has been written and then converted in MATLAB executable files. The simulation and testing has been performed using the 3-phase voltage source invter model shown in fig. 1 and fig. 2 shows the block diagram of modified switching scheme. Working of the modified scheme has been explained in fig. 8 as flowchart.

Fig.3 presents the reference signal for all three phase for time-period 'T'. This time-period 'T' is divided in six equal time intervals. The reference Ra is more than reference Ya, reference Ya is less than reference Ba and reference Ba is more than or equal to reference Ra, during the interval between t0 and t1. A similar comparison can be made for other time intervals. This comparison of reference helps in designing switching logic in such a way that, it can be applied to obtain output current of any value of frequency, without much change in programming. Instead of defining logic of switching in terms of time interval, the respective value of all the three references has been used for it. The same has used in designing the flow chart of programming.

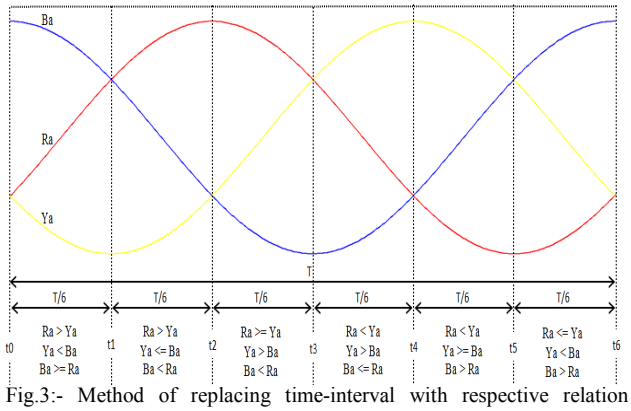


Fig.3:- Method of replacing time-interval with respective relation between references.

The IGBT switches ate controlled through the gate pulses generated from the switching circuitry. The ON pulses are represented by logic '1' and similarly OFF pulses by logic '0'. The description of the switching scheme is as follows.

Step 1) Initialize the circuit.
Step 2) Define parameters Tm (time-period), Ts (sampling time), t (set time) etc.
Step 3) Define initial reference signals as in (1).
Step 4) If$(Ra > Ya)\&\&(Ya < Ba)\&\&(Ba \geq Ra)$; then to step 5. Else jump to step 9.
Step 5) If$(Ra > Rb)$; then G1 = 1 & G4 = 0. Else G1 = 0 & G4 = 0.
Step 6) If$(Ya < Yb)$; then G3 = 0 & G6 = 1. Else G3 = 0 & G6 = 0.

Step 7) If($Ba > Bb$); then G5 = 1 & G2 = 0. Else G5 = 0 & G2 = 1.
Step 8) Jump to step 34.
Step 9) If($Ra > Ya$)&&($Ya \leq Ba$)&&($Ba < Ra$); then to step 10. Else jump to step 14.
Step 10) If($Ra > Rb$); then G1 = 1 & G4 = 0. Else G1 = 0 & G4 = 0.
Step 11) If($Ya < Yb$); then G3 = 0 & G6 = 1. Else G3 = 1 & G6 = 0.
Step 12) If($Ba < Bb$); then G5 = 0 & G2 = 1. Else G5 = 0 & G2 = 0.
Step 13) Jump to step 34.
Step 14) If($Ra \geq Ya$)&&($Ya > Ba$)&&($Ba < Ra$); then to step 15. Else jump to step 19.
Step 15) If($Ra > Rb$); then G1 = 1 & G4 = 0. Else G1 = 0 & G4 = 1.
Step 16) If($Ya > Yb$); then G3 = 0 & G6 = 1. Else G3 = 1 & G6 = 0.
Step 17) If($Ba < Bb$); then G5 = 0 & G2 = 1. Else G5 = 0 & G2 = 0.
Step 18) Jump to step 34.
Step 19) If($Ra < Ya$)&&($Ya > Ba$)&&($Ba \leq Ra$); then to step 20. Else jump to step 24.
Step 20) If($Ra < Rb$); then G1 = 0 & G4 = 1. Else G1 = 0 & G4 = 0.
Step 21) If($Ya > Yb$); then G3 = 1 & G6 = 0. Else G3 = 0 & G6 = 0.
Step 22) If($Ba < Bb$); then G5 = 0 & G2 = 1. Else G5 = 1 & G2 = 0.
Step 23) Jump to step 34.
Step 24) If($Ra < Ya$)&&($Ya \geq Ba$)&&($Ba > Ra$); then to step 25. Else jump to step 29.
Step 25) If($Ra < Rb$); then G1 = 0 & G4 = 1. Else G1 = 0 & G4 = 0.
Step 26) If($Ya > Yb$); then G3 = 1 & G6 = 0. Else G3 = 0 & G6 = 1.
Step 27) If($Ba > Bb$); then G5 = 1 & G2 = 0. Else G5 = 0 & G2 = 0.
Step 28) Jump to step 34.
Step 29) If($Ra \leq Ya$)&&($Ya < Ba$)&&($Ba > Ra$); then to step 30. Else jump to step 34.
Step 30) If($Ra < Rb$); then G1 = 0 & G4 = 1. Else G1 = 1 & G4 = 0.
Step 31) If($Ya < Yb$); then G3 = 0 & G6 = 1. Else G3 = 0 & G6 = 0.
Step 32) If($Ba > Bb$); then G5 = 1 & G2 = 0. Else G5 = 0 & G2 = 0.
Step 33) Jump to step 34.
Step 34) Update the reference signal
Step 35) Calculate t = t + Ts.
Step 36) If ($t \geq Tm$) ; then reset 't'. Else go to 36.
Step 37) Jump to 4.

IV. SIMULATION RESULTS

Simulation and the testing have been performed on the 3-phase voltage source inverter as shown in fig.1. Linear as well as nonlinear loading conditions have been used for testing of the developed scheme. The linear loading has been realized through R-L load while rectifier has been used for nonlinear loading condition. Along with developed switching scheme, PWM has also been tested and the results of both schemes have been compared for analytical study.

The simulation has been performed under the balanced loading condition for the output current of 50 Hz. The results and waveforms, presented in the paper are recorded under the steady state condition. The various parameters and loading conditions are as follows:

1) Source voltage (Vdc) = 700V
2) Source impedance (Zs) = 0.1 Ω
3) Modulating index for PWM = 0.8
4) Modulation frequency for PWM = 1.2KHz
5) R–L load = 10+ j23.5619 Ω
6) Rectifier load
   (a) Resistance = 15 Ω
   (b) Capacitance = 100 μf

A. *Linear* load

The Simulation and testing has been performed and observations have been made for all the three phases of the output but results and waveforms presented in paper are for R phase only as the model is operating under the balanced condition and all the three phase are giving similar outputs. Waveforms of voltage and current for 3-phase voltage source inverter using PWM scheme and Modified-switching scheme is shown in fig. 4 and fig. 5 respectively.

The results obtained from simulation of the model are in Table 1. The comparative analysis has been done between PWM and newly developed switching scheme.

We can observe that the new switching scheme has performed better than its counterpart PWM scheme under linear loading condition. The DC side voltage source is set as 700 V. The RMS value of output voltage for each phase obtained from new scheme is 289.9 V which is significantly much more than that obtained using PWM scheme. Also the value of RMS current 8.786 A, obtained from new switching scheme is more than the RMS current value obtained from PWM scheme. The new scheme also presents a significant improvement in THD

in voltage as well as in current. The THD in voltage obtained from PWM is 91.60% while same using new proposed switching scheme is 31.32%. Similarly, THD in output current obtained from PWM is 2.73% and same for new proposed switching scheme is 1.00%. In case of proposed scheme the DC components in voltage and current are 0.08102 V and 0.02939 A respectively. In case of PWM the DC components in voltage and current are 7.388 V and 0.7453 A respectively.

B. Nonlinear load

Nonlinear loading conditions for simulation and testing has been realized using rectifier with resistance. Output waveforms of voltage and current are shown in fig. 6 and fig. 7 for PWM and modified switching scheme respectively.

The results obtained by simulating the model are presented in Table 2. Observation and analysis has been done using this result.

Similar to linear load all the testing has been performed under the balanced three phase condition. The DC voltage source for the testing and simulation is taken as 700 V. The rectifier load has been connected to the inverter through and inductive coil. The values have been monitored between the output port of inverter and inductive coil. The results presented in the paper are for R phase only.

In this case also the RMS and peak as well as THD values of output voltage and current obtained from newly developed switching scheme show significant improvement over PWM scheme. RMS value of voltage is 212V in new proposed scheme as compared to 196.7V obtained from PWM scheme. The RMS value of current is 11.45 A for new proposed scheme which is more than 11.12A obtained from PWM scheme. THD in voltage waveform are 59.14% and 92.50% for new scheme and PWM scheme respectively. THDs in current waveform are 4.02% and 2.15% from PWM and modified scheme respectively. The new switching strategy also shows significant improvement in terms of DC component in voltage and current. The DC component in voltage and current are 5.832 and 0.5593 for PWM scheme respectively. The same is 0.9474 and 0.0314 in voltage and current waveforms obtained from modified switching scheme.

Table 1
Simulation results for 3 phase voltage source inverter with R-L (linear) load.

Parameters	PWM Scheme	Modified Switching Scheme
V rms	197.9	289.9
I rms	6.004	8.786
V peak	279.9	410
I peak	8.49	12.42
V THD	91.60%	31.32%
I THD	2.73%	1.00%
DC Component in V	7.388	0.08102
DC Component in I	0.7453	0.02939

Table 2
Simulation results for 3 phase voltage source inverter with Rectifier (nonlinear) load.

Parameters	PWM Scheme	Modified Switching Scheme
V rms	196.7	212
I rms	11.12	11.45
V peak	278.2	299.9
I peak	15.73	16.2
V THD	92.50%	59.14%
I THD	4.02%	2.15%
DC Component in V	5.832	0.5593
DC Component in I	0.9474	0.0314

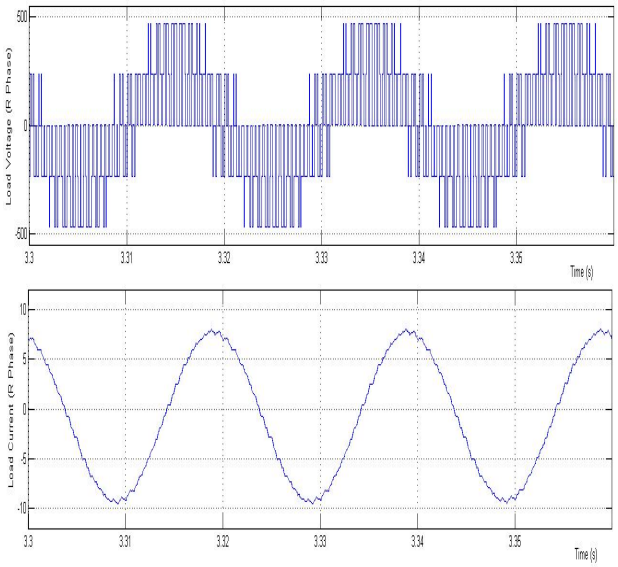


Fig.4 Three-phase VS DC–AC inverter output (load voltage and load current) waveform of R- Phase under linear loading using PWM scheme.

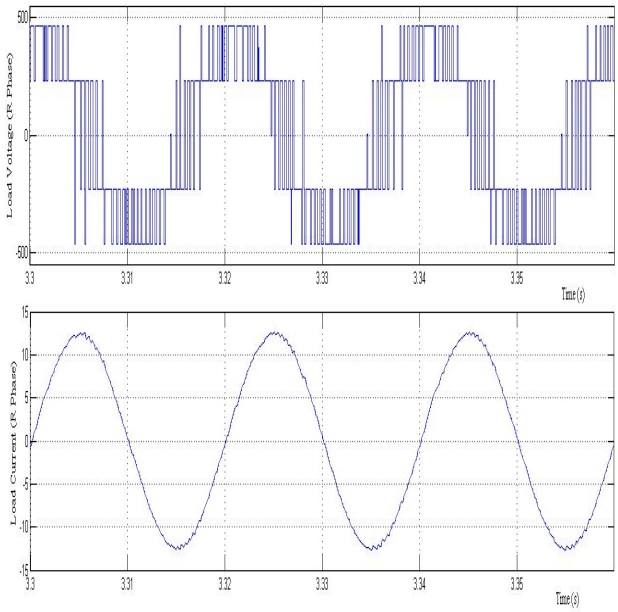


Fig.5 Three-phase VS DC–AC inverter output (load voltage and load current) waveforms under linear loading using modified switching scheme.

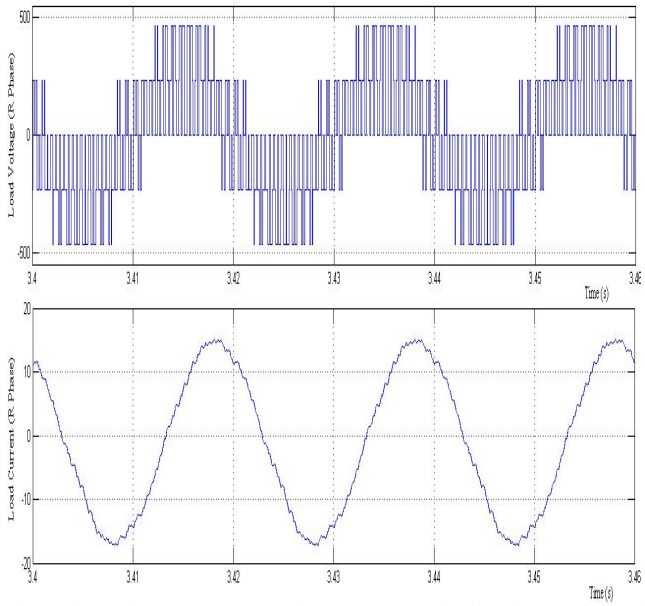


Fig.6 Three-phase VS DC–AC inverter output (load voltage and load current) waveform of R- Phase under non- linear loading using PWM scheme.

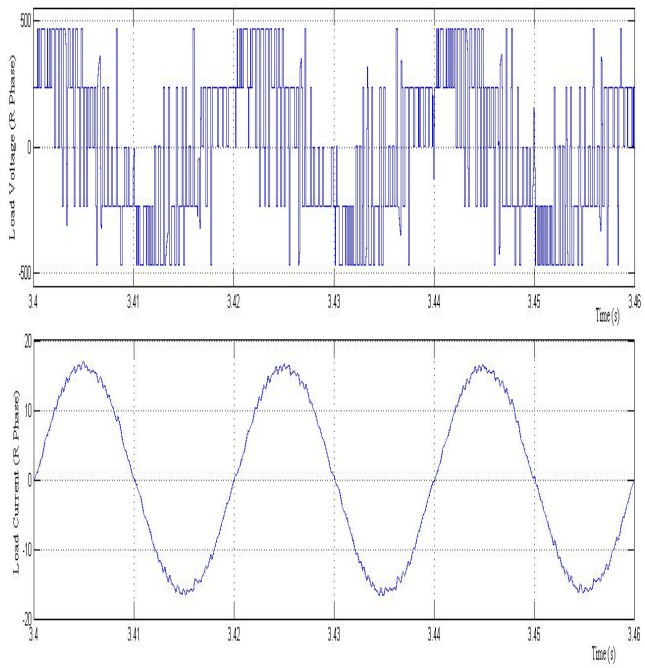


Fig.7 Three-phase VS DC–AC inverter output (load voltage and load current) waveform of R- Phase under non- linear loading using modified scheme

V. CONCLUSION

Modified switching scheme using for 3- phase voltage source inverter using value of load current as a modulating signal has been developed. The uniqueness of the new strategy lies in its reference signals where instead of having a fixed reference signal and scaling the modulating signal to that range, a new concept of variable reference signal has been introduced. Simulation and testing of new scheme have been performed on SIMULINK facility of MATLAB. Testing includes the linear as well as nonlinear loading condition for both scheme i.e. PWM and Modified scheme. Results obtained from both schemes have been compared in terms of RMS values, Peak values, THD and DC component for both voltage as well as current waveforms.

Both in the case of linear as well as non-linear loading condition, the newly developed scheme has shown significant improvement over the PWM scheme in terms of RMS and Peak values of current and voltage. In terms of THD as well DC component in output voltage and current waveforms the modified scheme presents improved results.

REFERENCES

[1]. S. A. Saleh, C. R. Moloney and M. A. Rahman, "Experimental Performance of the Single-Phase Wavelet-Modulated Inverter," IEEE Trans. Power Electronics, vol. 26, no. 9, pp. 2650 – 2600, Sep. 2011.

[2]. S. A. Saleh, C. R. Moloney and M. A. Rahman, "Developing a Non – Dyadaic MRAS for Switching DC – AC inverters," in Proc. IEEE 12th DSP Conf. , Jackson Lake Lodge, WY, Sep. 2006, pp. 544 – 549.

[3]. S. A. Saleh, C. R. Moloney and M. A. Rahman, "Development and Testing of Wavelet Modulation for Single-Phase Inverters," IEEE Trans. Ind. Electronics, vol. 56, no. 7, pp. 2588 – 2599, July 2009.

[4]. S. A. Saleh, "the Implementation and Performance Evaluation of 3 VS Wavelet Modulated AC – DC Converters," IEEE Trans. Power Electronics, vol. 28, no. 3, pp. 1096 – 1106, March 2013.

[5]. Pankaj H. Zope, Pravin G. Bhangale, Prashant Sonare and S. R. Suralkar, "Designing and Implementation of carrier based Sinusoidal PWM Inverter," in International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 1, Issue 4, October 2012, ISSN: 2278 – 8875.

[6]. Phoivos D. Zigas, "The Delta Modulation Technique in Static PWM Inverters" in IEEE Trans. on Industrial Applications, vol.IA-17, No. 2, March/April 1981, pp. 199-204.

[7]. M. Azizur Rahman, Jhon E. Quaicoe, and M. A. Choudhary, "Performance Analysis of Delta Modulated PWM Inverters," in IEEE Trans. on Power Electronics, Vol. PE-2, No. 3, July 1987, PP. 227 – 233.

[8]. B. Vasantha Reddy and B. Chitti Babu, "Hysteresis Controller and Delta Modulator – Two Viable Scheme for Current Controlled Voltage Source Inverter," in IEEE International Conference on Technical Postgraduates 2009, 14-15/Dec/2009, Kuala Lumpur, Malaysia.

[9]. Dr. P. S. Bimbhra, "Power Electronics," 4th ed. Khanna Publishers, Delhi:2009

[10]. Jai P. Agrawal, " Power electronic systems – Theory and Design," 1st ed. Pearson Education Asia, Singapore:2001

[11]. L. Umanand, "Power Electronics Essentials & Applications," 1st ed. Wiley, Delhi:2010

[12]. Ankit Yadav and M M Tripathi, "A Novel Switching Scheme for Single Phase VS Inverter Using Value of Output Current as Modulating Signal" in International Journal of Science & Technology, ISSN (online): 2250-141X, Vol. 3 Issue 3, December 2013.

[13]. Stephen J. Chapman, "MATLAB Programming for Engineers," 3rd ed. Thomson Learning, Singapore: 2005.

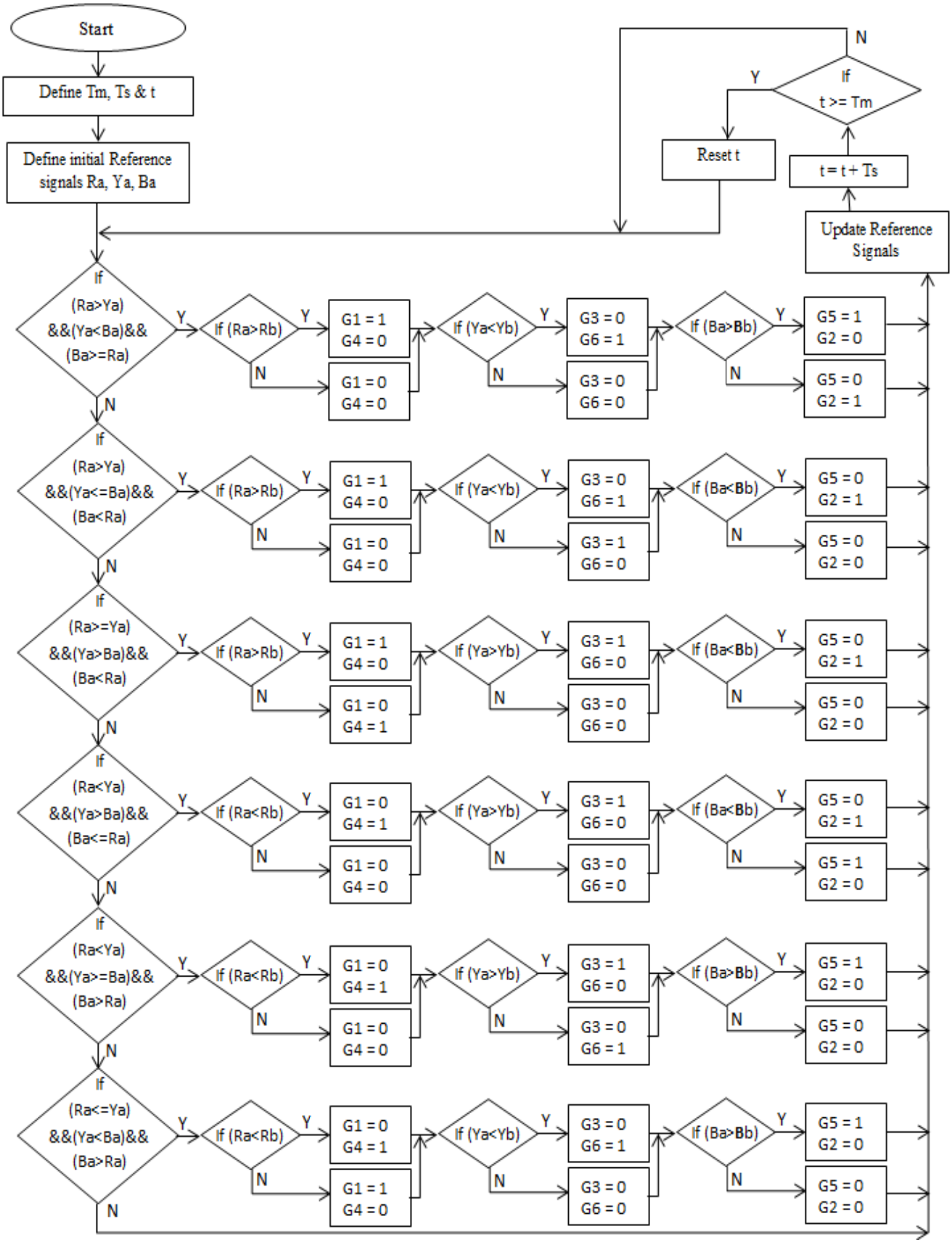


Fig.8- Flowchart for new switching scheme for 3-phase voltage source inverter

Nature & Environment
Vol. 20 (1), 2015: 23-32
Website: www.natureandenvironment.com

ISSN (Print) : 2321-810X
ISSN (Online) : 2321-8738

RESEARCH ARTICLE

Use of ecofriendly in Low Temperature Circuit in Terms of First Law and Second Law Efficiency of Three Stages Cascade Refrigeration of Biomedical Applications

R.S. Mishra
Mechanical Engineering Department,
Delhi Technological University Delhi-110042 (India)
Email: rsmishra1651956@yahoo.co.in

Received: 17th October 2014, Revised: 15th December 2014, Accepted: 22nd December 2014

ABSTRACT

Biomedical preservation requires storing biological specimens like stem cells blood and organs, at a storage temperature below -95°C, Hence, the main aim of the current research is to conduct a thermodynamic energy and exergy analysis to determine the thermal performance of three stages cascade refrigeration systems in the high temperature condenser temperature using R1234ze and R1234yf in high temperature circuit of condenser temperature of 70°C and zero degree centrigade of first evaporator temperature in High temperature circuit, and evaporator temperature(-80°C) in intermediate circuit using R134a and R410a as ecofriendly refrigerants is investigated. For low temperature evaporator temperature (-135°C) using hydrocarbons (R290, R600 and R600a), R404aand other refrigerants in low temperature evaporator circuit on system performances (i.e. overall system coefficient of performance, (first law efficiency), exergetic efficiency (second law efficiency) and system exergy destruction ratio (EDR) is investigated in three stages /four stages cascade refrigeration systems are shown in this paper

***Keywords**: Three stages cascade refrigeration system, Energy–Exergy analysis, Low temperature cascade system, Irreversibility Prediction, Biomedical applications.*

INTRODUCTION

Biomedical preservation requires storing biological specimens like stem cells,sperm, blood and organs, at a storage temperature below -95°C. For long-term storage of biological materials, temperatures below -100°C are generally considered to safeguard against the effects of denitrification and Crystallization (ASHRAE Handbook, 2006). The use of a single-cycle vapour compression refrigeration system can only achieve effective cooling of about -40°C, and the efficiency begins to deteriorate under -35°C due to the vast difference between the evaporating and condensing temperatures. Thus, in order to reach a lower temperature, a cascade refrigeration system is utilized NIST (2007); Nasruddin and Syaka (2012). In cascade refrigeration systems consist of at least two refrigeration systems that work independently. The two or more than refrigeration systems are connected by a cascade heat exchanger, where heat is released in the condenser low-temperature circuit (LTC) and is absorbed from the evaporator high-temperature circuit (HTC).The hydrocarbon refrigerants have good thermo physical properties and are environmentally friendly can be used for low temperature applications below -80°C (COX N. 2007).

Hence, the main aim of the current research is to conduct a thermodynamic energy and exergy analysis to determine the thermal performance of three stages / four stages cascade refrigeration systems by varying the high temperature condenser temperature from 0°C to 70°C using R1234ze (GWP =6) and R1234yf (GWP=4) in high temperature circuit with zero degree centigrade evaporator temperature in High temperature circuit, and (-80°C) evaporator temperature in intermediate circuit using R134a and R410a as ecofriendly refrigerants is investigated. The effect of various ecofriendly refrigerants in the low temperature evaporator temperature (-135°C) using hydrocarbons (R290, R600 and R600a), R404a and other refrigerants in low temperature evaporator circuit is also investigated.

LITERATURE REVIEW

Several researchers such as Cox *et.al* (2007), Dopazo *et. al.* (2009), Kim *et. al.* (2008), Lee *et. al.* (2006), Nasruddin and Syaka (2012), Rahadiyan (2007) have evaluated the thermodynamic performance of the two-stage cascade refrigeration systems. Bhattacharyya et al. (2005) studied a carbon dioxide–propane (R744–R290) optimum cascade evaporating system to define an evaporating temperature of R744 for application in heating circuits. Getu and Bansal(2008) analyzed a carbon dioxide–ammonia (R744–R717) cascade system thermodynamically to determine the optimum condensing temperature of R744 in the low-temperature circuit and mass flow ratio, to give the system maximum COP in terms of sub-cooling, superheating, evaporating temperature, condensing temperature and temperature difference in the system's cascade condenser. The thermodynamic analysis of the carbon dioxide–ammonia (R744–R717) cascade system by Dopazo Alberto, *et al.* (2009) employed both exergy analysis and energy optimization, to determine the optimum condensing temperature of R744 in the low-temperature circuit. Lee et al. (2006) also analyzed a carbon dioxide–ammonia (R744–R717) cascade system thermodynamically to determine the optimum condensing temperature of R744 in the low-temperature circuit.

RESEARCH GAPS IDENTIFIED

Although lot of research work available on the two stages cascade refrigeration systems for lower temperature of circuit up to minimum temperature up to (-70^{o}C) but using ecofriendly new refrigerants such as R1234ze (of GWP six and zero ODP) and R1234yf (of GWP four and zero ODP) in three stages cascade refrigeration systems maintaining temperature below -95 ^{o}C for biomedical applications etc is not reported in literature so far. Thepresent investigation mainly deals with energy –exergy analysis of three stage cascase vapour compression refrigeration system using 70^{o}C condenser temperature and zero degree centigrade evaporator temperature and -80^{o}C of second evaporator and -135^{o}C of low temperature evaporator circuit was investigated. Similarly the effect of temperature overlapping in the cascade systems is also known as approach1 (Cascade condenser1-evaporator temperature1) and approach 2 (cascade condenser temperature-evaporator temperature2) and approach 3 (cascade condenser temperature3-evaporator temperature3) on system performances (i.e. overall system coefficient of performance, (first law efficiency), exergetic efficiency (second law efficiency) and system exergy destruction ratio (EDR) is also investigated in three stages cascade refrigeration systems. The results of present investigated systems are shown in this paper

SYSTEM DESCRIPTION

The systems chosen for present investigation is three stages cascade vapour compression refrigeration system, which consist of R1234ze /R1234yf in high temperature evaporator circuit (HTC) using condenser temperature of 70^{o}C while R134a/R410a in the intermediate temperature cascade evaporator circuit (ICT) and hydro-carbons and other ecofriendly refrigerants in the low temperature (-135^{o}C) evaporator circuit (LTC). The performance equations using energy exergy analysis for finding irreversibilities in the system as well as components the three stages cascade vapour compression refrigeration system have been formulated and performance equations have been analysed by using EES and results for above system have been shown in Table-1 to 6 respectively.

RESULTS AND DISCUSSIONS

The performance of three stages cascade refrigeration system using R1234ze in hot temperature evaporator circuit and R1234a in intermediate temperature circuit and six ecofriendly refrigerants in low temperature evaporator circuit are shown in presented Table-1(a) to Table 1(d) respectively. It was observed that using R1234ze in high temperature circuit gives better performance than using ecofriendly R1234yf refrigerant. It was also observed that using R134a in the intermediate temperature circuit gives better thermal performance in terms of overall system COP and exergetic efficiency (in terms of

second law efficiency)than using R410a in the intermediate temperature circuit. The exergy destruction ratio is larger by using R1234yf in high temperature circuit while using R410a in the intermediate temperature circuit is also higher than using R134a in the intermediate temperature circuit.

Table 1(a): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond} = 70°C, $T_{cascade\ _eva1}$ = 0°C, $T_{cascade_eva2}$ = -80°C, T_{eva3} = -135°C, Temperature overlapping (Approach1) = Approach2 = 10(°C) and $Compressor_{Efficiency1}$ = 0.80, $Compressor_{Efficiency2}$ = 0.80, $Compressor_{Efficiency3}$ = 80%.

Ecofriendly Refrigerant in LTC	System COP	System EDR	Exergetic Efficiency
R404a	0.2230	2.867	0.2586
Ethylene	0.2179	2.958	0.2526
Ethane	0.2241	2.848	0.2599
R290	0.2266	2.806	0.2627
R600a	0.2306	2.741	0.2673
R600	0.2299	2.752	0.2665

Table 1(b): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature Temperature -overlapping (Approach1) =Approach2 = 10 (°C) and $Compressor_{Efficiency1}$ = 0.80, $Compressor_{Efficiency2}$ = 0.80 $Compressor_{Efficiency3}$= 0.80

Refrigerant in LTC	System COP	System EDR	Exergetic Efficiency
R404a	0.2081	3.145	0.2413
Ethylene	0.2034	3.241	0.2358
Ethane	0.2091	3.125	0.2424
R290	0.2114	3.080	0.2451
R600a	0.2150	3.011	0.2493
R600	0.2144	3.023	0.2486

Table 1(c): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C ,T_{eva3}= - 135 °C, Temperature overlapping (Approach1)=Approach2= 10(°C) and $Compressor_{Efficiency1}$=0.80 $Compressor_{Efficiency2}$= 0.80, $Compressor_{Efficiency3}$= 80%

Ecofriendly Refrigerant in LTC	System COP	System EDR	Exergetic Efficiency
R404a	0.2228	2.871	0.2583
Ethylene	0.2177	2.962	0.2524
Ethane	0.2239	2.852	0.2596
R290	0.2264	2.810	0.2625
R600a	0.2303	2.745	0.2670
R600	0.2297	2.756	0.2663

Table 1(d): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Refrigerant in LTC	System COP	System EDR	Exergetic Efficiency
R404a	0.2079	3.149	0.2410
Ethylene	0.2032	3.245	0.2355
Ethane	0.2089	3.129	0.2422
R290	0.2114	3.084	0.2448
R600a	0.2148	3.015	0.2491
R600	0.2142	3.027	0.2483

Similarly the circuit first law efficiency COP of low temperature circuit) is significantly affecting using different ecofriendly refrigerants in the low temperature evaporator circuit used for biomedical applications. The Circuit COP using R600a is better than using other ecofriendly refrigerants. Therefore by using safety measures one can use R600a in the three stages cascade refrigeration system due to better first law and second law performances.

Table 2(a): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Refrigerant in LTC	COP_{HTC}	COP_{ITC}	COP_{LTC}
R404a	1.797	1.118	1.163
Ethylene	1.797	1.118	1.117
Ethane	1.797	1.118	1.173
R290	1.797	1.118	1.196
R600a	1.797	1.118	1.234
R600	1.797	1.118	1.228

Table 2(b): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Refrigerant in LTC	COP_{HTC}	COP_{ITC}	COP_{LTC}
R404a	1.544	1.118	1.163
Ethylene	1.544	1.118	1.117
Ethane	1.544	1.118	1.173
R290	1.544	1.118	1.196
R600a	1.544	1.118	1.234
R600	1.544	1.118	1.228

Table 2(c): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Refrigerant in LTC	COP_{HTC}	COP_{ITC}	COP_{LTC}
R404a	1.797	1.116	1.163
Ethylene	1.797	1.116	1.117
Ethane	1.797	1.116	1.173
R290	1.797	1.116	1.196
R600a	1.797	1.116	1.234
R600	1.797	1.116	1.228

Table 2(d): First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Refrigerant in LTC	COP_{HTC}	COP_{ITC}	COP_{LTC}
R404a	1.544	1.116	1.163
Ethylene	1.544	1.116	1.117
Ethane	1.544	1.116	1.173
R290	1.544	1.116	1.196
R600a	1.544	1.116	1.234
R600	1.544	1.116	1.228

As approach-1 (temperature overlapping between low temperature condenser and intermediate evaporator circuit) decreases, the first law efficiency (overall system COP) increases and also second law efficiency in terms of exergetic efficiency increases.

Table 3(a): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	System COP	System EDR	Exergetic Efficiency
15	0.2110	3.087	0.2447
10	0.2230	2.867	0.2586
5	0.2352	2.667	0.2727
0	0.2477	2.483	0.2871

Table 3(b): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	System COP	System EDR	Exergetic Efficiency
15	0.1970	3.378	0.2284
10	0.2081	3.145	0.2413
5	0.2193	2.932	0.2543
0	0.2303	2.738	0.2675

Table 3(c): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	System COP	System EDR	Exergetic Efficiency
15	0.2109	3.090	0.2445
10	0.2228	2.871	0.2583
5	0.2339	2.672	0.2723
0	0.2472	2.489	0.2866

Table 3(d): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	System COP	System EDR	Exergetic Efficiency
15	0.1969	3.381	0.2283
10	0.2079	3.149	0.2410
5	0.2190	2.938	0.2539
0	0.2304	2.744	0.2671

As approach-2 (temperature overlapping between low temperature condenser and intermediate evaporator circuit) decreases, the first law efficiency (overall system COP) increases and also second law efficiency in terms of exergetic efficiency increases.

Table 4(a): Effect of Approach2 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C,

$T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature Temperature - overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80 $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	System COP	System EDR	Exergetic Efficiency
15	0.2108	3.092	0.2444
10	0.2230	2.867	0.2586
5	0.2360	2.654	0.2737
0	0.2499	2.451	0.2898

Table 4(b): Table-4(b) Effect of Approach2 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	System COP	System EDR	Exergetic Efficiency
15	0.1968	3.383	0.2282
10	0.2081	3.145	0.2413
5	0.2201	2.919	0.2552
0	0.2328	2.704	0.270

4(c): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	System COP	System EDR	Exergetic Efficiency
15	0.2106	3.092	0.2441
10	0.2228	2.871	0.2583
5	0.2358	2.658	0.2734
0	0.2497	2.455	0.2895

Table 4(d): Effect of Approach1 on First law performance (overall system) Exergetic Efficiency (second law efficiency) and system exergy Destruction Ratio (EDR) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	System COP	System EDR	Exergetic Efficiency
15	0.1966	3.387	0.2279
10	0.2079	3.149	0.2410
5	0.2198	2.923	0.2549
0	0.2326	2.708	0.2697

Table 5(a): Effect of Approach1 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80, $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.797	1.018	1.163
10	1.797	1.118	1.163
5	1.797	1.228	1.163
0	1.797	1.351	1.163

Table 5(b): Effect of Approach1 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80 $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.544	1.018	1.163
10	1.544	1.118	1.163
5	1.544	1.228	1.163
0	1.544	1.351	1.163

Table 5(c): Effect of Approach1 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.797	1.017	1.163
10	1.797	1.116	1.163
5	1.797	1.225	1.163
0	1.797	1.346	1.163

Table 5(d): Effect of Approach1 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80 $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-1	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.544	1.017	1.163
10	1.544	1.116	1.163
5	1.544	1.225	1.163
0	1.544	1.346	1.163

Table6(a): Effect of approach2 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80 $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.797	1.118	1.055
10	1.797	1.118	1.163
5	1.797	1.118	1.289
0	1.797	1.118	1.436

Table 6(b): Effect of approach2 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R134a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80, $Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.544	1.118	1.055
10	1.544	1.118	1.163
5	1.544	1.118	1.289
0	1.544	1.118	1.436

Table 6(c): Effect of approach2 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234ze refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature -overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.797	1.116	1.055
10	1.797	1.116	1.163
5	1.797	1.116	1.289
0	1.797	1.116	1.436

Table 6(d): Effect of approach2 on First law performance various circuit COPs) of Four stages cascade refrigeration system using ecofriendly R1234yf refrigerant in the high temperature evaporator circuit and ecofriendly R410a in intermediate cascade evaporator temperature circuit and six ecofriendly Refrigerants used in low temp evaporator circuit for T_{cond}= 70 °C, $T_{cascade\ _eva1}$= 0 °C, $T_{cascade_eva2}$= - 80 °C , T_{eva3}= - 135 °C, Temperature-overlapping (Approach1) =Approach2= 10 (°C) and $Compressor_{Efficiency1}$= 0.80,,$Compressor_{Efficiency2}$= 0.80 $Compressor_{Efficiency3=}$ 0.80

Effect of Approach-2	COP_{HTC}	COP_{ITC}	COP_{LTC}
15	1.544	1.116	1.055
10	1.544	1.116	1.163
5	1.544	1.116	1.289
0	1.544	1.116	1.436

CONCLUSION

The following conclusions have been drawn from present investigations-

i.The use of R1234ze refrigerant has GWP =6) gives better thermal performance than using R1234yf (GWP=4) in the high temperature circuit although both refrigerants have zero ozone depletion potential.

ii. By using R410a in the intermediate temperature circuit the thermal performance decreases and system exergy destruction (EDR) is increases as compared with using R134a in the intermediate temperature circuit.

iii.By using safety measures, the use of R600a hydrocarbon gives better thermal performances than other hydrocarbons such as R290 and R600 and other ecofriendly refrigerants (R404a, ethane and ethylene etc).

REFERENCES

1. ASHRAE (2006): Handbook Refrigeration System and Applications, American Society of Heating, Refrigerating, and Air-Conditioning Engineer, Inc., Atlanta, Georgia.
2. Bhattacharyya S, Mukhopadhyay S, Kumar A, Khurana RK and Sarkar J. (2005): Optimization of a CO2–C3H8 cascade system for refrigeration and heating. International Journal of Refrigeration; Vol-28: 1284–1292.
3. Cox N. (2007): Working towards more environmentally friendly Refrigerant Blends. 12th European Conference, Milano, Italy, June 8 – 9.
4. Dopazo JA, Fernandez-Seara J, Sieres J and Uhi FJ. (2009): Theoretical Analysis of a CO2-NH3 Cascade Refrigeration System for Cooling Applications at Low Temperatures, Applied thermal engineering; Vol-29: 1577-1583.
5. Gettu HM and Bansal PK. (2008): Thermodynamic Analysis of an R744-R717 Cascade Refrigeration system. International Journal of Refrigeration; 31: 45-54.
6. Kim JH, Cho JM and Kim MS. (2008): Cooling performance of several CO2/propane mixtures and glide matching with secondary heat transfer fluid. International Journal of Refrigeration; Vol-31: 800-806.
7. Lee TS, Liu CH and Chen TW. (2006): Thermodynamic Analysis of Optimal Condensing Temperature of Cascade-Condenser In CO2/NH3 Cascade Refrigeration Systems. International Journal Of Refrigeration; Vol-29 :1100-1108
8. Nasruddin and Syaka DRB (2013): Thermodynamics Analysis of Refrigerant Selection in Cascade Refrigeration System, International Conference of Saving Energy in Refrigeration and Air-Conditioning. Departemen Teknik Mesin Fakultas Teknik-Universitas Indonesia,
9. Niu, Boulian and Zhang, Yufeng. (2007): Experimental Study of the Refrigeration Cycle Performance for R744/R290 Mixtures. International Journal of Refrigeration; Vol-30:pp.37-42.
10. Rahadiyan L. (2007): Study of Propane and Ethane Characteristics in Cascade Refrigeration System, Thesis, Department of Mechanical and Precision Engineering. The Graduate School of Gyeongsang National University, Gyeongsang.

Nature & Environment
Vol. 20 (1), 2015: 37-44
Website: www.natureandenvironment.com

ISSN (Print) : 2321-810X
ISSN (Online) : 2321-8738

RESEARCH ARTICLE

Sustainability analysis of eight ecofriendly refrigerants two-stage vapour compression refrigeration system-using energy-exergy performance

Kapil Chopra[1], V. Sahni[2] and R.S. Mishra[3]

[1]Mechanical Engineering, M.V.D.V., Katra (Jammu)
[2]Mechanical Engineering, S.L.I.T., Longowal (Punjab)
[3]Department of Mechanical, Production & Industrial and Automobiles Engineering
Delhi Technological University Delhi-110042
Email: rsmishra1651956@yahoo.co.in

Received: 17th October 2014, Revised: 21st December 2014, Accepted: 30th December 2014

ABSTRACT

In this paper, comparative analysis of R152a,R600,R600a,R410a,R290,R1234yf,R404a and R134a as refrigerants in two stage vapour compression refrigeration system has been done on the basis of energetic and exergetic performance. Performance parameters such as entropy generations, COP, exergetic efficiency, sustainability index were investigated at different ambient condition. It was found that both energy and exergy efficiencies of R134a is 8.97% and 5.38% lower than R152a and R600 respectively. It was also observed that Irreversibility was minimal at higher evaporating temperatures and condenser was responsible for highest irreversibility or losses in two stage vapour compression refrigeration system. Sustainability index for R152a (1.96) was highest compared to other refrigerants.

Keywords*: COP, Irreversibility, Exergetic efficiency*

NOMENCLATURE

CFC	chlorofluorocarbon	Evap	evaporator
HCFC	hydrochlorofluorocarbon	Cond	condenser
Q	rate of heat transfer (kW)	amb	ambient conditions
W	work (kW)	WI	water intercooler
T	temperature (^oC)	H	high temperature
s	specific entropy (kJ/kgK)	L	low temperature
h	specific enthalpy(kJ/kg)	LP	low pressure
$\dot{S}$	entropy (kJ/s.K)	EV	expansion valve
ODP	ozone depletion potential	gen	generation
GWP	global warming potential	HP	high pressure
m	mass flow rate (kg/s)	FC	flash chamber
η	efficiency (non-dimensional)	ex	exergetic

INTRODUCTION

Refrigeration technology based on the principle of rejection of heat to the surrounding at higher temperature and absorption of heat at low temperature (i.e. Kapil Chopra, V.Sahni, R.S Mishra. (2014)), evaporator, expansion valve, condenser and compressor are the main four components of single stage vapour compression system. Vapour compression refrigeration systems consume large amount of electricity. This difficulty can be removed by improve the performance parameters of system. Coefficient of performance and exergetic efficiency are main two parameters to calculate the performance of refrigeration systems. Coefficient of performance can be enhanced either by minimizing power consumption of compressor or increasing of refrigeration effect. Refrigeration effect can be increased by adoption of multi-stage throttling .On the other hand power consumption of compressor can be enhanced by

incorporation of multi-stage compression and flash chamber. Collective effect of these two factors improves overall performance of vapour compression system.

It is presented that irreversibility in system components take place due to large temperature difference between system and surrounding. In order to improve the system performance Irreversibility should be measured in the cycle because Exergy losses are responsible for degradation of system performance .Coefficient of performance is commonly used to calculate the performance of vapour compression system but COP provides no information regarding thermodynamic losses in the system components. Using exergy analysis one can be quantify the exergy losses in vapour compression refrigeration systems. Exergy losses increase with increasing of temperature difference between systems and surrounding. Exergy is the available or useful energy and loss of energy means loss of exergy in the system. Exergy losses are useful to improve the performance of system and better utilization of energy input given to the system which is beneficial for environmental conditions and economics of energy technologies. Utilization of green energy can be increased by this method (i.e. Oktay, Z., and I. Dincer. 2007 Rosen, M.A., I. Dincer, and M. Kanoglu. 2008 Genoud, S., and J.B. Lesourd. 2009) In past decades, refrigerants such as R12,R02,R22 etc. used in vapour compression refrigeration system responsible for increasing of global warming and ozone depletion potential. An international society named Montreal protocol discussed and signed on the refrigerants having higher global warming and ozone depletion potential values for all countries. In order to control the emission of green house gases one more committee was formed named as Kyto protocol(E. Johnson, (1998)) .After 90's a program was ran to phase out the higher GWP and ODP refrigerants(CFC and HCFC) for the purpose of environmental problems..

To replace "old" refrigerants with "new" refrigerants, lots of researchers (i.e. E. Johnson, (1998), M. Padilla, R. Revellin and J. Bonjour, (2010) , H. O. Spauschus, (1988) , J. U. Ahamed, R. Saidur and H. H. Masjuki, (2011) ,R. Llopis, E. Torrella, R. Cabello and D. S_anchez, (2010) , A. Arora and S. C. Kaushik, (2008) , V. Havelsky, (2000)) have been carried out R. Selladurai and V. Saravanakumar[2013] evaluated thermal performance parameters such as COP and exergetic efficiency with R290/R600 hydrocarbon mixture on a domestic refrigerator designed to work with R134a and observed that performance of same system is higher with R290/R600a hydrocarbon mixture compared to R134a. In their analysis condenser, expansion valve and evaporator showing lower exergy destruction compared to compressor. V. S. Reddy, N. L. Panwar and S. C. Kaushik, (2012) presented a theoretical analysis of R134a, R143a, R152a, R404A, R410A, R502 and R507A in vapour compression refrigeration system and effect on coefficient of performance and second law efficiency with variation of superheating of evaporator outlet, evaporator temperature and degree of sub-cooling at condenser outlet, vapour liquid heat exchanger effectiveness and degree of condenser temperature was discussed. They reported that COP and exergetic efficiency significantly affected with change of evaporator and condenser temperatures and also observed that R134a and R407C show highest and lowest performance in all respect. S. Kumar, M. Prevost and R. Bugarel, (1989) carried out energy and exergy analysis of single stage vapour compression refrigeration system using R11 and R12 as working fluids. Thermal performance evaluation in terms of COP, exergetic efficiency and exergy losses in different components (compressor, evaporator, expansion valve and condenser) was done. Cornelissen, R.L. (1997) proposed that non-renewable energy sources are useful for minimizing the irreversibility of the system for sustainable development of systems. He also observed that emissions of gases put adverse effect on environmental conditions. Netherlands.C. Nikolaidis and D. Probert, (1998) considered the effect of condenser and evaporator temperatures on two-stage vapour compression refrigeration system using R22 was studied and suggested that there is requirement to optimize the condenser and evaporator conditions.

Many researchers carried out researches on different proportion of hydrocarbons as working fluid in vapour compression refrigeration systems. Fatouh, M., and E.I.M. Kafafy. (2006) suggested replacing R134a with hydrocarbon mixtures such as propane,

propane/isobutane/n-butane mixtures, butane, and various propane mass fractions in domestic refrigerator. Pure butane showed high operating pressures and low coefficient of performance among considered refrigerants. Wongwises, S., A. Kamboon, and B. Orachon. (2006) did experimental investigation on automotive air-conditioners with isobutene, propane, butane and suggested to replace R134a with these hydrocarbon mixtures. They observed that mixture of propane 50%, butane 40%, and isobutene 10% was best hydrocarbon mixture to replace R134a. Jung, D., C.B. Kim, K. Song, and B. Park. (2000)., Arcaklioglu, E. (2004) and Arcaklioglu, E., A. Cavosuglu, and A. Erisen. (2005). suggested to use of pure hydrocarbon instead of their mixtures due variation in condenser and evaporator temperature during phase changing at constant pressure. These Changes in condenser and evaporator temperature cause for problem in vapour compression refrigeration cycle. Liedenfrost, W., K.H. Lee, and K.H. Korenic. (1980] investigated Freon as refrigerant on the performance of a refrigeration cycle. Yumrutas, R., M. Kunduz, and M. Kanoglu. 2002 . observed the effect on exergy loss with change of evaporation and condenser temperature. Khan, S.H. (1992) studied that due to the low expansion process and compressor efficiency most of the irreversible losses occurred in the system.

Through above literature, it was found that energy, exergy and sustainable analysis of single stage vapour compression refrigeration systems have been done. But no literature contributed for energy and exergy analysis of two-stage vapour compression refrigeration system. Present works analyze the system in terms of energy and exergy efficiencies and explain the effect of exergy losses on two-stage vapour compression refrigeration system with hydrocarbons and R134a.

MATHEMATICAL CALCULATIONS

Some mathematical calculations are required to analyze the two-stage vapour compression refrigeration system based on energy and exergy method. Two stage vapour compression refrigeration system consist of low and high pressure compressor, condenser, evaporator, expansion valves, water-intercooler and flash chamber. Energy and exergy efficiencies are different for different ecofriendly refrigerants for same system. Following assumptions are taken for thermodynamic analysis of the system:

1. Temperature and pressure losses are not considered.
2. All components are running under steady state conditions.
3. Energy and exergy losses due to potential and kinetic energy are neglected.
4. Mechanical efficiencies of low and high pressure compressors are assumed to be 80%.

Two stage vapour compression refrigeration system and its P-H plot shown in Figure 1 and Figure 2 respectively. Exergy, energy and sustainability analysis can be done as follow:

Fig.1: Schematic diagram of two-stage vapour compression refrigeration system

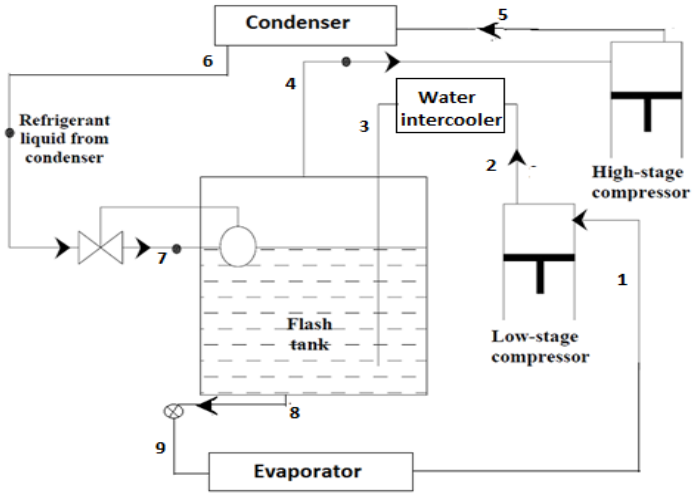
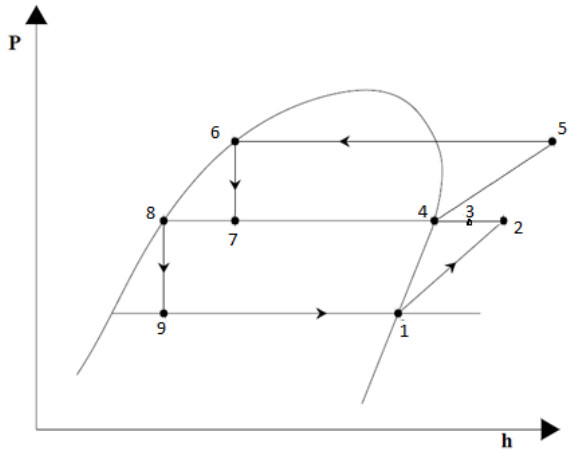


Fig. 2: Pressure enthalpy diagram of two-stage vapour compression refrigeration system



Exergy at any point, $EX = (h - h_{amb}) - T_{amb}(s - s_{amb})$

For high and low temperature compressors

$$(T_{amb}\dot{S}_{gen})_{HP} = \dot{W}_{HP} + \dot{m}_{HP}((h_1 - h_2) - T_{amb}(s_1 - s_2))$$

$$(T_{amb}\dot{S}_{gen})_{LP} = \dot{W}_{LP} + \dot{m}_{LP}((h_4 - h_5) - T_{amb}(s_4 - s_5))$$

For evaporator

$$(T_{amb}\dot{S}_{gen})_{Evap} = \dot{m}_{LP}((h_9 - h_1) - T_{amb}(s_9 - s_1)) + Q\left(1 - \frac{T_{amb}}{T_L}\right)$$

For condenser

$$(T_{amb}\dot{S}_{gen})_{Cond} = \dot{m}_{HP}((h_5 - h_6) - T_{amb}(s_5 - s_6)) - Q_{Cond}\left(1 - \frac{T_{amb}}{T_H}\right)$$

For expansion valves

$$(T_{amb}\dot{S}_{gen})_{EV} = \dot{m}_{HP}((h_6 - h_7) - T_{amb}(s_6 - s_7)) + \dot{m}_{LP}((h_8 - h_9) - T_{amb}(s_8 - s_9))$$

For water-intercooler

$$(T_{amb}\dot{S}_{gen})_{WI} = \dot{m}_{LP}((h_2 - h_3) - T_{amb}(s_2 - s_3))$$

For flash chamber

$$(T_{amb}\dot{S}_{gen})_{FC} = \dot{m}_{LP}((h_3 - h_4) - T_{amb}(s_3 - s_4)) + (\dot{m}_{HP}h_7 - \dot{m}_{LP}h_8) - T_{amb}(\dot{m}_{HP}s_7 - \dot{m}_{LP}s_8)$$

Exergetic efficiency

$$\eta_{ex} = \frac{Q\left(1 - \frac{T_{amb}}{T_L}\right)}{W_{LP} + W_{HP}}$$

Coefficient of Performance (COP)

$$COP = \frac{\dot{m}_{LP}((h_1 - h_9)}{W_{LP} + W_{HP}}$$

Sustainability index (SI)

$$SI = \frac{1}{1 - \eta_{ex}}$$

Total entropy generation

$$(T_{amb}\dot{S}_{gen})_{TOTAL} = \Sigma(T_o\dot{S}_{gen})_{HP} + (T_o\dot{S}_{gen})_{LP} + (T_o\dot{S}_{gen})_{Evap} + (T_o\dot{S}_{gen})_{Cond} + (T_o\dot{S}_{gen})_{EV} + (T_o\dot{S}_{gen})_{WI} + (T_o\dot{S}_{gen})_{FC}$$

RESULTS AND DISCUSSION

In this discussion effect of change of evaporator and condenser temperature on performance parameters like coefficient of performance, exergy loss, exergetic efficiency and sustainability index was studied for considered refrigerants.

CHANGE OF COEFFICIENT OF PERFORMANCE WITH CHANGE IN EVAPORATING TEMPERATURE FOR CONSIDERED REFRIGERANTS

As cleared from Table 1 that coefficient of performance of R134a is 4.2-8.9% and 3.6-5.3% lower than R152a and R600 respectively or in other words R134a consumes more electricity than R152a and R600.Ambient condition play an important role in electricity consumption of vapour compression refrigeration systems because higher the temperature difference between system and surrounding higher will be compressor work that's why COP of vapour compression refrigeration system increase with increase in evaporator temperature and decrease with decrease in evaporator temperature.COP of considered system with R152a, R600, R134a, R600a, R290, R410a, R1234yf and R404a varied between 1.45-4.88, 1.40-4.86, 1.33-4.69, 1.33-4.72, 1.30-4.57, 1.31-4.43, 1.30-4.47 and 1.12-4.15 respectively between -50 ^{o}C to 5 ^{o}C evaporator temperature.

Table 1: Variation of COP at different evaporating temperature for different refrigerants change of exergy loss with change in evaporating temperature for considered refrigerants

T_{Evap} (^{o}C)	R152a	R600	R134a	R600a	R410a	R290	R1234yf	R404a
-50	1.457	1.409	1.337	1.33	1.314	1.303	1.197	1.123
-45	1.59	1.54	1.466	1.458	1.437	1.429	1.321	1.24
-40	1.74	1.689	1.611	1.603	1.576	1.571	1.461	1.371
-35	1.911	1.857	1.775	1.767	1.732	1.733	1.62	1.521
-30	2.105	2.05	1.964	1.956	1.91	1.917	1.802	1.691
-25	2.328	2.273	2.181	2.175	2.115	2.129	2.013	1.887
-20	2.587	2.533	2.433	2.429	2.351	2.376	2.259	2.115
-15	2.89	2.838	2.73	2.729	2.627	2.666	2.548	2.383
-10	3.251	3.202	3.083	3.086	2.955	3.01	2.894	2.702
-5	3.686	3.641	3.509	3.519	3.348	3.425	3.312	3.086
0	4.219	4.182	4.033	4.052	3.831	3.936	3.827	3.558
5	4.889	4.863	4.691	4.723	4.436	4.578	4.476	4.151

As shown in Fig.3 exergy destructions or exergy losses decreases with increase of evaporator temperature. This is because that if evaporating temperature decreases the heat exchange between working fluid entered into the evaporator tubes and space being cooled also decreases, which finally decrease the cooling effect and therefore exergy destruction increases. Among selected refrigerants R404a (6.02-30.83 KW) and R152a (4.94-22.06 KW) shows higher and lower exergy loss for selected evaporator temperature range respectively. It was also observed that flash chamber, compressor, condenser, expansion valve, water-intercooler and evaporator are in increasing order of exergy loss for different refrigerants.

Fig. 3: Variation of exergy loss (KW) at different evaporating temperature for different refrigerants Change of exergy loss with change in evaporating temperature for R152a as working fluid

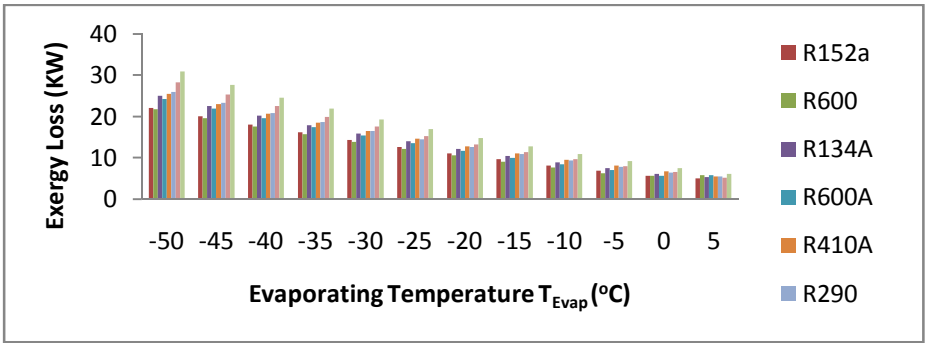


Fig.4 shows the variation of exergy loss for individual component with change in evaporating temperature with R152a used as working fluid. Behaviors of exergy destruction in different components of two stage vapour compression refrigeration system for rest of refrigerants are also observed similar. Flash chamber responsible for highest and evaporator shows lowest exergy destruction compared to other components. The exergy destruction in the components increase with the decrease of evaporating temperature. It was found that with increase in difference between evaporator and condenser temperatures exergy losses increases with R12, R134a, R22, and R502 used as refrigerants.

Fig. 4: Variation of exergy loss (KW) at different evaporating temperatures for R152a *Change of exergy loss with change in condensing temperature for considered refrigerants*

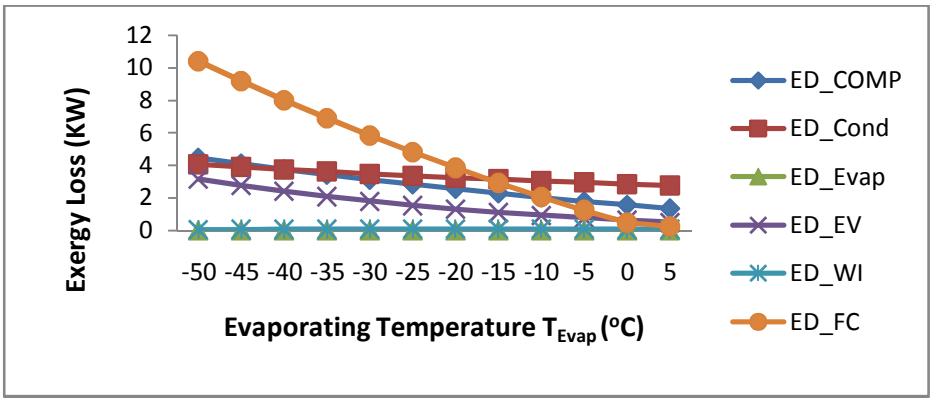
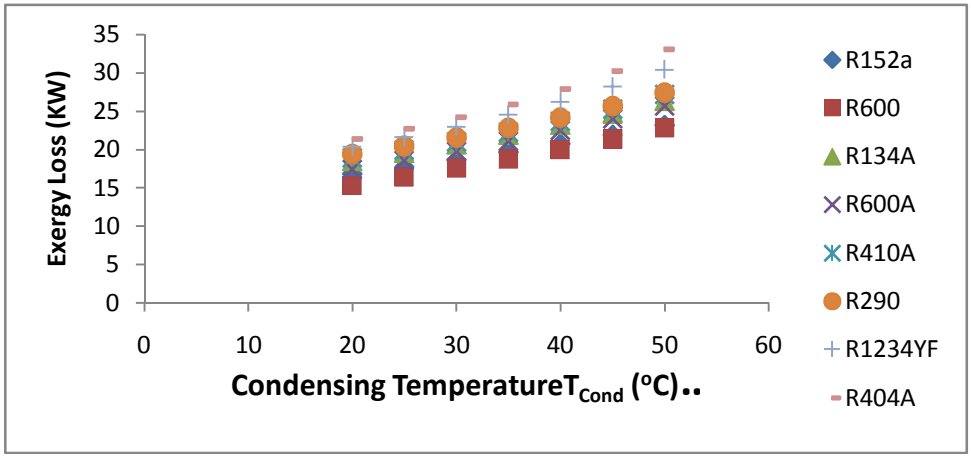


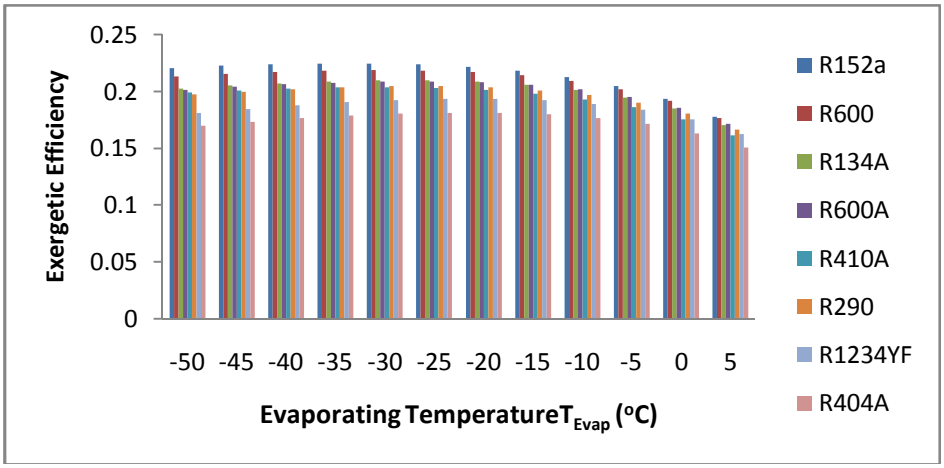
Fig. 5: Variation of exergy loss (KW) at different condensing temperatures for different refrigerants *Change of exergetic efficiency with change in evaporating temperature for considered refrigerants*



It is observed from Fig.5 that for all considered refrigerants exergy destructions increased with increase of condensing temperature. This is due to increase of temperature difference between condenser and surrounding.

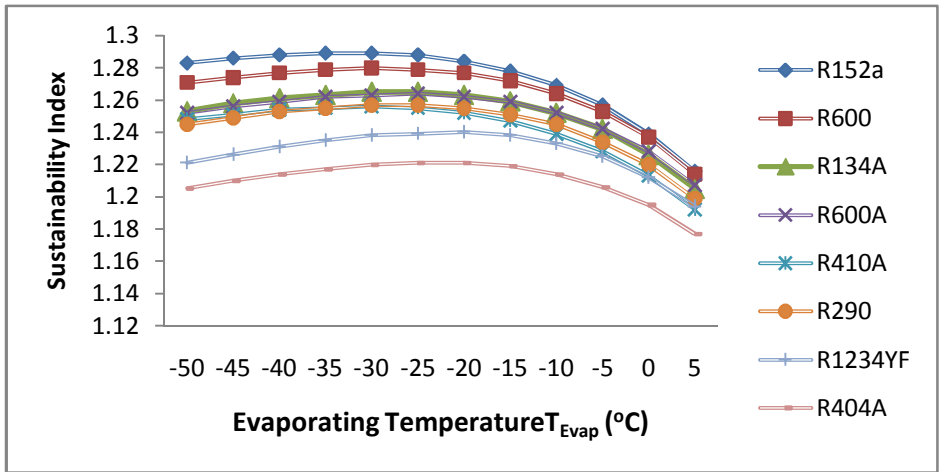
It is found that exergy losses decreases with increase of evaporating temperature for considered refrigerants.Fig.6 shows that R152a gives highest exergetic efficiency among selected refrigerants. The purpose of condenser to take out the heat produced by compressor in discharge line and carried by refrigerant during cooling effect in evaporator. This heat in refrigerant removed by transferring heat to the wall of condenser tubes due to convection and then transfer of heat due to conduction from tubes wall to surrounding.

Fig. 6: Variation of exergetic efficiency at different evaporating temperatures for different refrigerants *Variation of sustainability index with change in evaporating temperature for considered refrigerants*



As shown in Fig.7 with increase in evaporator temperature sustainability index increases for selected refrigerants. R152a shows higher sustainability index than R134a for selected evaporating temperature range. It is also found that R152a and R600 have higher sustainability index and low impact on surrounding compared to R134a.

Fig. 7: Variation of sustainability index at different evaporating temperatures for different refrigerants



CONCLUSIONS AND RECOMMENDATION

Energetic and exergetic analysis of two stage refrigeration system was carried out with different refrigerants and following conclusion and recommendation are presented below-

1. R404a shows lowest performance among selected refrigerants.
2. Exergy destruction for R134a is higher than R152a and R600 but lower than R600a, R410a, R290, R1234yf and R404a.
3. Exergetic and energetic efficiency of R152a is highest among selected refrigerants.
4. Flash chamber responsible for highest exergy destruction for all refrigerants taken under consideration.
5. Sustainability index of the R152a and R600 are higher than that of R134a at every evaporator temperature. It indicates also less environmental impact for hydrocarbons.

However the performance of R152a and R600 is higher than R134a but hydrocarbons are flammable in nature and can be used in limited applications. Therefore R134a recommended for all kind of applications.

REFERENCES

1. Arora and S. C. Kaushik, (2008): Theoretical analysis of a vapour compression refrigeration system with R502, R404A and R507A, Int. J. Refrig. 31 998–1005.
2. Arcaklioglu, E. (2004). Performance comparison of CFCs with their substitutes using artificial neural networks. International Journal of Energy Research 28 (12): 1113–25.
3. Arcaklioglu, E., A. Cavosuglu, and A. Erisen. (2005). An algorithmic approach towards finding better refrigerant substitutes of CFCs in terms of the second law of thermodynamics. Energy Conversion and Management 46: 1595–1611.
4. Cornelissen, R.L. (1997). Thermodynamics and sustainable development. Ph.D. thesis. University of Twenty, the Netherlands.
5. C. Nikolaidis and D. Probert, (1998) Exergy method analysis of a two-stage vapour-compression refrigeration plants performance, Appl. Energy 60 241–256.
6. E. Johnson, (1998) Global warming from HFC, Environ.Impact Asses. 18, pp.485–492.
7. Fatouh, M., and E.I.M. Kafafy. (2006). Assessment of propane/commercial butane mixtures as possible alternatives to R134a in domestic refrigerators. Energy Conversion and Management 47:2644–58..
8. Genoud, S., and J.B. Lesourd. (2009). Characterization of sustainable development indicators for various power generation technologies. International Journal of Green Energy 6 (3): 257–267
9. H. O. Spauschus, (1988) HFC 134a as a substitute refrigerant for CFC 12, Int. J. Refrig. 11, pp. 389–392.
10. J. U. Ahamed, R. Saidur and H. H. Masjuki, (2011). A review on exergy analysis of vapor compression refrigeration system, Renew. Sustain. Energy Rev. 15, pp.1593–1600.
11. Jung, D., C.B. Kim, K. Song, and B. Park. (2000). Testing of propane, iso-butane mixture in domestic refrigerants. International Journal of Refrigeration 23: 517–27.
12. Kapil Chopra, V.Sahni, R.S Mishra. (2014) Thermodynamic analyses of multiple evaporators vapour compression refrigeration systems with R410A, R290, R1234YF, R502, R404A and R152A.International Journal of Air-conditioning and Refrigeration 21(1) pp. 1-14. Khan, S.H. 1992. Second law based thermodynamics analysis of vapor compression system. M.S.Thesis in Engineering, Department of Mechanical Engineering, King Fahad University of Petrolium and Minerals, Saudi Arabia.
13. Liedenfrost, W., K.H. Lee, and K.H. Korenic. (1980). Conversion of energy estimated by second law analysis of power consuming process. Energy 5: 47–61.
14. M. Padilla, R. Revellin and J. Bonjour, (2010) Exergy analysis of R413A as replacement of R12 in a domestic refrigeration system, Energy Convers.Manag. 51, pp.2195–2201.
15. Oktay, Z., and I. Dincer. (2007). Energetic, exergetic, economic and environmental assessments of the bigadic geothermal district heating system as a potential green solution. International Journal of Green Energy 4 (5): 549–569.
16. Rosen, M.A., I. Dincer, and M. Kanoglu. (2008). Role of exergy in increasing efficiency and
17. R. Llopis, E. Torrella, R. Cabello and D. S_anchez, (2010) Performance evaluation of R404A and R507A refrigerant mixtures in an experimental double-stage of vapour compression plant, Appl. Energy 871546–1553.
18. R. Saravanakumar and V. Selladurai (2013)., Exergy analysis of a domestic refrigerator using eco-friendly R290/R600a refrigerant mixture as an alternative to R134a, J. Therm. Anal. Calorim.
19. S. Kumar, M. Prevost and R. Bugarel, (1989) Exergy analysis of a vapour compression refrigeration system,Heat Recovery Syst. CHP 9 ,pp.151–157.
20. Yumrutas, R., M. Kunduz, and M. Kanoglu. (2002). Exergy analysis of vapor compression refrigeration systems. Exergy, an International Journal 2: 266–72.
21. V. Havelsky, (2000) Investigation of refrigerating system with R12 refrigerant replacements, Appl. Therm.Eng. 20, 133–140.
22. V. S. Reddy, N. L. Panwar and S. C. Kaushik, (2012) Exergy analysis of a vapour compression refrigeration system with R134a, R143a, R152a, R404A, R407C, R410A, R502 and R507A, Clean Techn.Environ. Policy 14 47–53.
23. Wongwises, S., A. Kamboon, and B. Orachon. (2006). Experimental investigation of hydrocarbon mixtures to replace HFC-134a in an automotive air conditioning system. Energy Conversion and Management 47: 1644–59.

Microsyst Technol
DOI 10.1007/s00542-015-2552-1

TECHNICAL PAPER

Novel back-contact back-junction SiGe (BC-BJ SiGe) solar cell for improved power conversion efficiency

Rahul Pandey[1] · **Rishu Chaujar**[1]

Received: 2 November 2014 / Accepted: 18 April 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In this paper, a SiGe based Back-Contact Back-Junction (BC-BJ) device structure called BC-BJ SiGe solar cell has been proposed. Photo reflection is significantly reduced in UV/Visible spectrum region in case of SiC/Si_3N_4/SiO_2 passivated BC-BJ SiGe solar cell. Result, indicates that presence of SiC play an important role in photoelectric conversion. Ray tracing and finite difference time domain (FDTD) algorithms are used to simulate optoelectronics characteristics of the device. Simulation achieves the barrier height of 0.8 eV for holes at the interface which results in a higher field. The lower interface recombination rate of the order of 10^{17} cm^{-3} s^{-1} has been obtained. The device shows improved photovoltaic parameters. External quantum efficiency >84 % in the spectrum range of 450–700 nm wavelength and more than 80 % in the range of 350–700 nm wavelength is obtained. Further, we obtained the fill-factor (FF) and power conversion efficiency (PCE), 79 %, 17.8 % and 79 %, 14.8 %, using FDTD and ray tracing methods, respectively. All the simulations have been done using atlas and devedit device simulator.

1 Introduction

Over the last 53 years, commercial silicon photovoltaics (PV) have been developed to convert sunlight into electricity and provide the most feasible carbon–neutral route to displacing terawatts(TW) of nonrenewable power consumed worldwide (Garnett and Yang 2010). Since solar cell technology is not an individual function of efficiency but also depends on the cost, hence large scale implementation is not economically feasible. In order to maximise solar cell efficiency, it is necessary to optimize both the electrical device characteristics and the optical absorption of the device (Deceglie et al. 2012; Deinega et al. 2013; Tian et al. 2007). Power conversion efficiency of Si-solar cell remains below 10.5 % for thin devices (Green et al. 2014). However, it has been observed that reducing the thickness of the photovoltaic cell to sub 10 μm can potentially have considerable impact on cost reduction.

One promising high-efficiency solar cell concept is the back-contact back- junction solar cell (Kerschaver and Beaucarne 2006) having both, junction and the electrodes at back side of the device. The first 17 % efficient under 50 sun condition BC-BJ solar cell presented by (Schwartz and Lammert 1975). The conventional BC-BJ solar cell have large area which result in higher material cost. Microsystem enabled 14 μm thick, 250 μm wide rear contact solar cell, with 14.9 % efficiency are presented by (Cruz-Campa et al. 2011), work shows that the ultrathin wafer handling problem can be avoided by using low cost microsystem techniques. Further, an all back-contact ultra-thin i.e., 10 μm thick silicon nanocone solar cell with 13.7 % power conversion efficiency is reported (Jeong et al. 2013). Further, back-contacted silicon heterojunction solar cells with 9 cm^2 active cell area were fabricated on 250-μm-thick, n-type, 4-in float-zone (FZ) wafers, with a nominal resistivity of 3 Ω cm, resulting in 21.5 % PCE (Tomasi et al. 2014). Also, independently confirmed conversion efficiency of 23.0 % for n-type rear contact cells has been reported (Feldmann et al. 2014) and at

✉ Rahul Pandey
rahulpandey@dtu.ac.in; rahulp.electronics@gmail.com

Rishu Chaujar
rishu.phy@dce.edu

[1] Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

Published online: 01 May 2015

the end of the year, the interdigitated back contact (IBC) solar cells developed which resulted in an independently confirmed designated-area efficiency of 24.4 %, The cell was fabricated on a 230 µm thick 1.5 Ω cm n-type CZ wafer (Franklin et al. 2014). These cells have higher thickness they require high quality materials, in order to have higher minority carrier lifetime: order of magnitude 10 ms. These kind of c-Si are very expensive, resulting in increased module cost.

Here, we propose a 10 µm thin novel BC-BJ SiGe solar cell with a 14.8 % power conversion efficiency that will overcome the problems associated with thin devices such as auger and surface recombination. Here, SiGe is used because the energy of the visible part of the spectrum used by solar cell is less than 50 % and in order to increase the absorption of sunlight, low bandgap (<1.1 eV) material is required. Using low bandgap material in solar cell makes the conversion efficiency of the cell effective. Nowadays, SiGe is used for high frequency transistors as well as potential has been explored for solar cells also. SiGe has higher carrier lifetime and higher mobility as compared to silicon. In SiGe technology, the electrical properties of silicon are modified with germanium. SiGe processing is simple because the physical and electrical properties of silicon and germanium are similar (Ouellete 2002). Also SiGe possesses higher mechanical strength, and suppresses the high intensity degradation of solar cell under illumination (Yang et al. 2010). The band gap of SiGe can be modified to vary by changing the Ge ratio (Wang et al. 2012; Povolny et al. 2000) and alternatively by varying the optical absorption, absorption is directly related to band gap of semiconductor

2 Device structure: parameters and simulation models

The simulation device structure i.e., BC-BJ SiGe solar is shown in Fig. 1. Device contains intrinsic SiGe wafer having length in y-direction equal to 10 µm and in x-direction equals to 100 µm, without optical absorber coating on top. The $Si_{0.9}$ $Ge_{0.1}$ was used during the simulation, very small fraction of germanium is used. Because, if we increase the germanium content the bandgap of SiGe decreases (Atlas user manual 2014), which further results in lower open circuit voltage. The width of $N+/P+$ regions are 30 and 45 µm respectively, in X direction whereas the depth is 100 and 300 nm in Y direction. The pitch of the device is 52 µm. The device consists of ohmic contacts and the doping concentration of $N+/P+$ regions are 4×10^{20} cm^{-3}. All simulations have been performed using atlas and devedit device simulator. The complex refractive index of

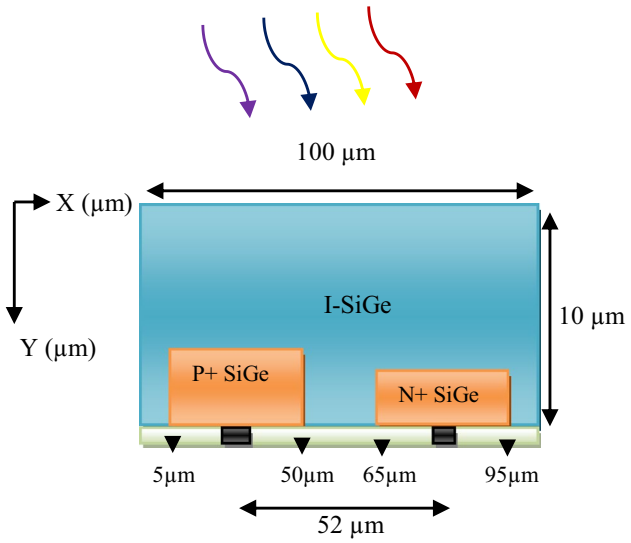


Fig. 1 Proposed SiGe based device: BC-BJ SiGe solar cell

Table 1 Parameters of SiGe used in simulation

Energy band gap (Eg)	1.08 eV
Band density (Nc)	2.8×10^{19} (cm^{-3})
Band density (Nv)	1.04×10^{19} (cm^{-3})
Intrinsic carrier conv (Ni)	1.45×10^{10} (cm^{-3})
Minority carrier lifetime for electron	3×10^{-5} s
Minority carrier lifetime for holes	1×10^{-5} s
Electron mobility (μ_n)	1430 cm^2/Vs
Hole mobility (μ_p)	480 cm^2/Vs

various materials as a function of temperature and composition has been obtained from sopra database. SiGe parameters are obtained from Silvaco library and shown in Table 1.

The energy band diagrams from $P+$–I–$N+$ regions are examined under dark and illuminated conditions, respectively. It is clearly observed that electron quasi Fermi level lies within the valence band and conduction band at $P+$ region and $N+$ region, respectively shown in Fig. 2a. This helps in analyzing the motion of the carriers i.e., if an electron is generated in the I region, then it finds build-in potential at $P+$ side and hence, will not travel towards $P+$ side but as there is no barrier in $N+$ side, it can easily move toward $N+$ region. In Fig. 2b, under illumination, the electron quasi Fermi level and band diagram is changed because of the light generated carriers. In intrinsic SiGe region, as the number of free carriers are increased during illumination, Fermi level moves closer to the bottom of the conduction band. The $N+$, $P+$ region acts as a back surface field (BSF) for holes and electrons, respectively. Thereby reflecting, the hole

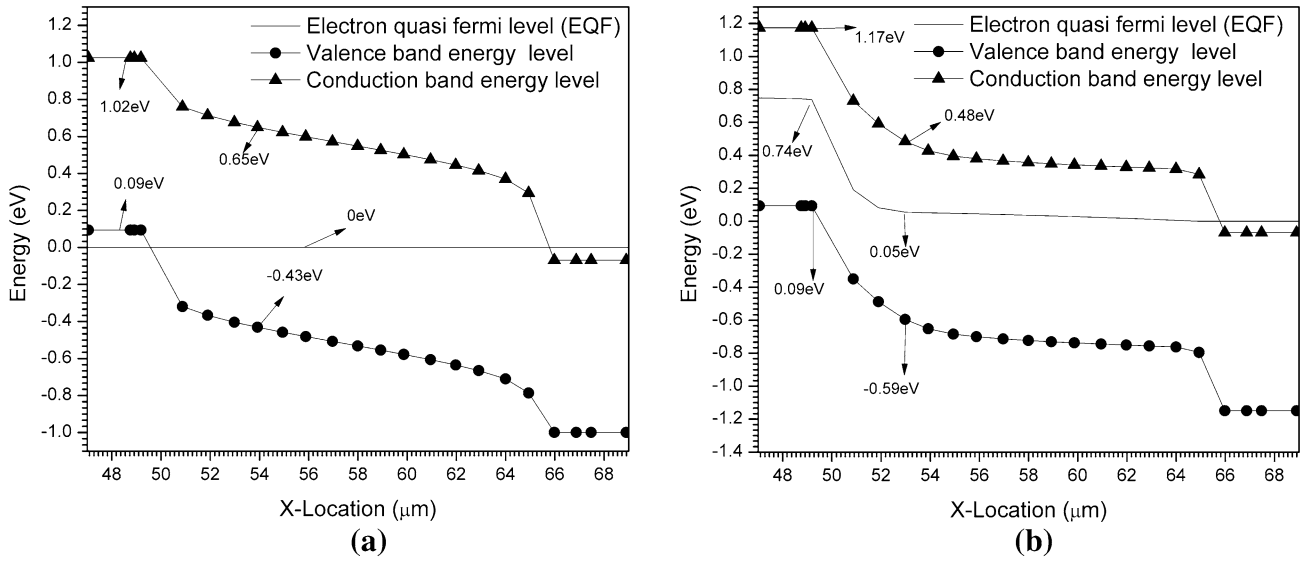


Fig. 2 The energy band diagrams from *P*+–*I*–*N*+ regions with Electron Quasi Fermi Level at zero bias V = 0. **a** Under dark and **b** With Illumination Pin = 0.1 W/cm^2

moving towards *N*+ region and electron moving towards *P*+ region by the barrier at *I*-*N*+ region and *I*-*P*+ region, respectively.

Further, in the presence of heavy doping (>10^{18} cm^{-3}), experimental work has shown that the pn product in silicon turn out to be doping dependent (Slotboom 1977). As the doping level rises, a decrease in the bandgap occurs, where the valence band raises by approximately same amount as the conduction band is lowered. Proposed device has heavily doped *P*+ and *N*+ regions and so bandgap narrowing effects are enabled. These effects may be described by an analytic expression relating the variation in bandgap (Eg) to the doping concentration, N. The expression used in simulation is from Slotboom and de Graaf (Slotboom and de Graaff 1976). Furthermore, Auger and Shockley- Read- Hall (SRH) recombination models are selected during the simulations. Auger recombination occurs through a three particle transition whereby a mobile carrier is either captured or emitted. The underlying physics for such processes is unclear and normally a more qualitative understanding is sufficient (Dziewior and Schmid 1977; Selberherr 1984). In SRH, Phonon transitions occur in the presence of a trap or defect within the forbidden gap of the semiconductor (Shockley and Read 1952; Hall 1952). The Poisson equation together with the continuity equation for electrons and holes are solved simultaneously by the software to obtain the current density- voltage (J-V) characteristics. The standard air mass 1.5 (AM1.5) spectrum is used.

3 Simulation results

3.1 Designing of highly absorbing surface passivation layer (HAPL) using carbide, oxide and nitride

In this section, we have designed and discussed the HAPL layer for rear contact SiGe solar cell, using carbide, and oxide and nitride layers. We have simulated a (10 × 10) μm thick intrinsic SiGe wafer and with different absorbing layers. Spectrum response i.e., reflectivity and available photo current as a function of wavelength is presented in Fig. 3a, b with different absorbing layers for intrinsic SiGe wafer. Results, indicate significant decrease in photo reflectivity in UV region due to presence of SiC. However, results also indicate SiC is highly reflective in visible region, therefore only SiC layered device shows lower available photocurrent in visible spectrum shown in Fig. 3b. So, in order to avoid this problem, oxide and nitride coated layers has been placed on SiC. Result shows, SiO_2 coated SiC provide better absorption compared to Si_3N_4 coated SiC shown in Fig. 3a, b. Result also indicates SiGe wafer without absorbing layer is highly reflective. Further, highest absorption is achieved in SiC/Si_3N_4/SiO_2 coated SiGe compared to other layers. Furthermore, results also indicate that SiC/SiO_2 coated SiGe wafer shows approximately same response as SiC/Si_3N_4/SiO_2 coated SiGe. This indicates SiC/SiO_2 coated SiGe can be used instead of SiC/Si_3N_4/SiO_2 coated SiGe in order to avoid complexity.

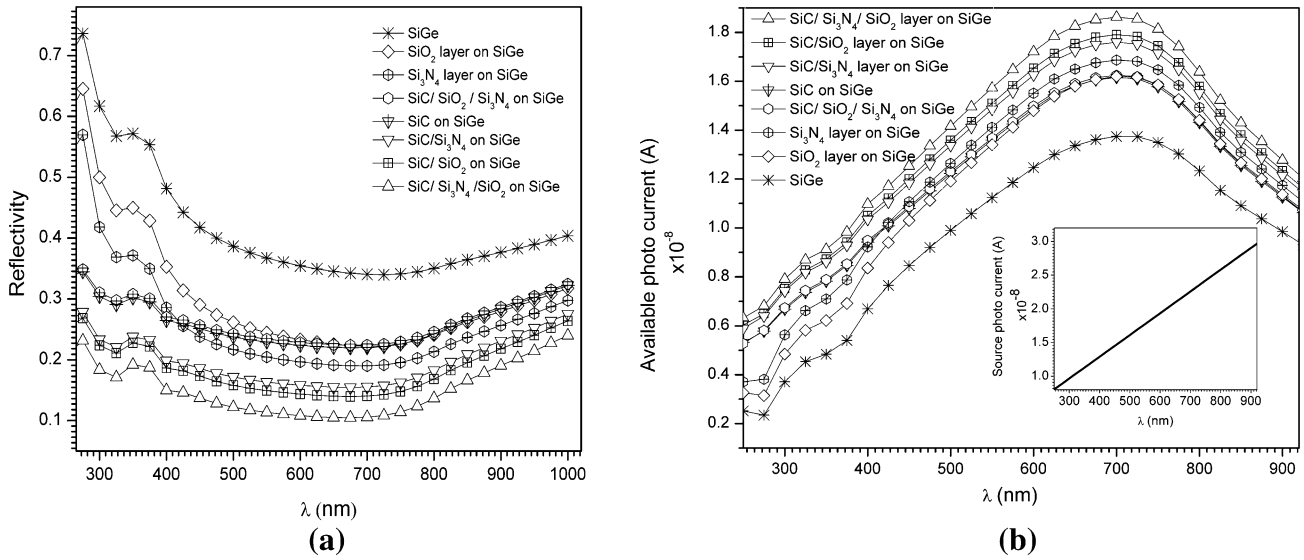


Fig. 3 Spectrum response of (10 × 10) μm thick *I-SiGe* wafer with different absorbing layers. **a** Reflectivity plot, **b** available photocurrent. *Inset* shows source photocurrent

3.2 Optical analysis of $SiC/Si_3N_4/SiO_2$ coated SiGe solar cell

In previous section, result shows that $SiC/Si_3N_4/SiO_2$ coated *I-SiGe* wafer shows superior spectrum response compared to other layers. In this section, we have performed $SiC/Si_3N_4/SiO_2$ coating on proposed device discussed earlier in device section. Thickness of $SiC/Si_3N_4/SiO_2$ layers are 0.05/0.05/0.1 μm respectively, simulated modified structure is shown in Fig. 4. The spectrum

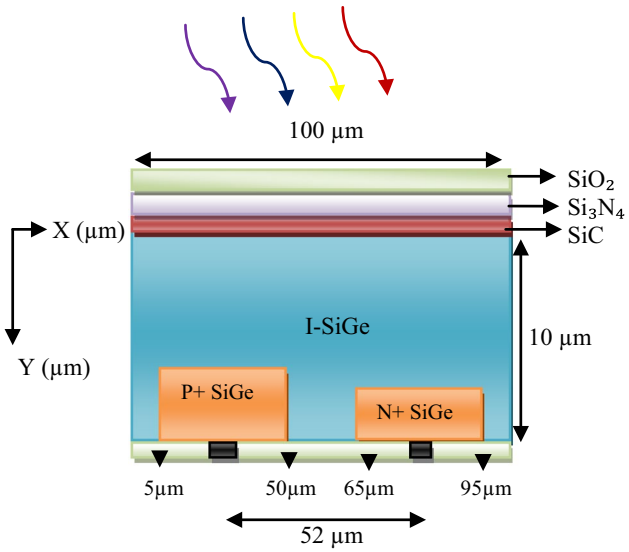


Fig. 4 Proposed SiGe based device: BC-BJ SiGe solar cell with $SiC/Si_3N_4/SiO_2$ coating on top

response of (BC-BJ) SiGe solar cell with conventional Si_3N_4 and $SiC/Si_3N_4/SiO_2$ layers on top is obtained and presented in Fig. 5a, b. From Fig. 5a, it is clear that an optical absorption >85 % is achieved in the spectrum range of 400–700 nm wavelength which higher than the conventional Si_3N_4 passivated SiGe cells. Result shows, significant improvement in absorption as well as EQE of $SiC/Si_3N_4/SiO_2$ passivated solar cell. Significant improvement has been observed in UV/Visible spectrum in $SiC/Si_3N_4/SiO_2$ passivated SiGe solar cell compared to conventional Si_3N_4 passivated cell, thereby increasing the photon absorption of proposed device. Higher absorption results in higher EQE shown in Fig. 4b. The device achieved EQE >80 % in the wavelength range of 400–700 nm. At wavelength equivalent to 600 nm it shows EQE of the order of 87.6 %. Also, in UV region, at wavelength equivalent to 250 nm $SiC/Si_3N_4/SiO_2$ passivated device shows 77 % EQE, whereas Si_3N_4 passivated device shows 43 % EQE. This results in 79 % higher EQE in $SiC/Si_3N_4/SiO_2$ passivated device compared to Si_3N_4 cell, at 250 nm wavelength.

3.3 Analysis of $SiC/Si_3N_4/SiO_2$ coated SiGe solar cell parameters: surface recombination rate and carrier concentrations

Further, in solar cell technology an absorbing layer should also provide good surface passivation, which prevents the recombination of the carriers at the interface. Our previous work shows that the presence of SiC as front surface passivation layer for rear contact silicon solar cell results in improved photovoltaic parameters compared to

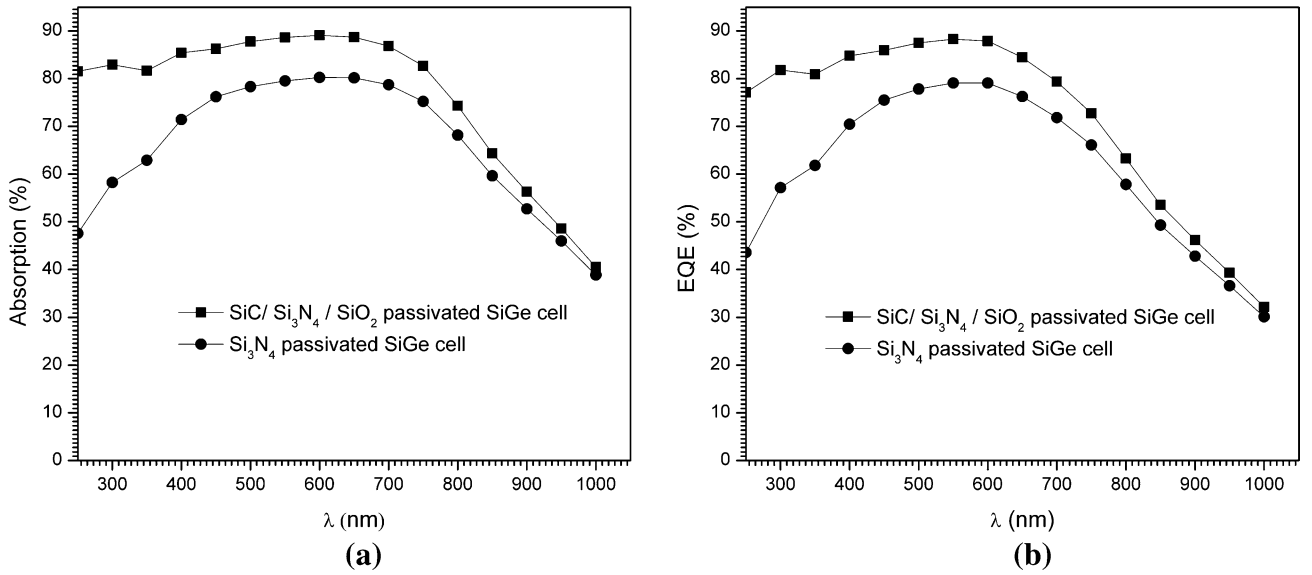


Fig. 5 Comparison of optical parameters of $SiC/Si_3N_4/SiO_2$ and Si_3N_4 passivated device: **a** absorption spectrum of BC-BJ SiGe Solar Cell. **b** External quantum efficiency of BC-BJ SiGe Solar Cell

conventional front surface passivation scheme (Pandey and Chaujar 2014a, b). Here, we have studied the $SiC/Si_3N_4/SiO_2$ based front surface passivation on BC-BJ SiGe solar cell. We have obtained the holes quasi Fermi energy level near the *SiC/I-SiGe* interface. Results are compared with BC-BJ SiGe solar cell having only Si_3N_4 as a front surface passivation. The obtained results are plotted in Fig. 6a–c. Result shows, in the proposed device higher field at the interface is achieved. We obtained the barrier height of ~0.8 eV for holes, in proposed device with $SiC/Si_3N_4/SiO_2$ based front surface passivation. Higher field will assist the holes to move towards the junction, resulting in low concentration of carriers (holes) at the interface shown in Fig. 6c. After illumination, at the interface, the carrier concentration of 2.15×10^9 cm^{-3} and 7.4×10^8 cm^{-3} for electron and holes, respectively has been observed in the device, whereas the bulk concentration of the carrier is ~10^{15} cm^{-3}. Further, Fig. 6b shows the comparison of interface recombination rate. In surface recombination process, an electron from conduction band recombines with the hole in the valance band via a defect level within the bandgap (Aberle 2010). If the concentration of the carriers is drastically reduced, then recombination rate reduces strongly. Further, Results reveal that the magnitude of interface recombination rate is less in the proposed device. The magnitude of interface recombination rate of order of 10^{17} cm^3 s^{-1} has been achieved in proposed device shown in primary axis in Fig. 6b. This section concludes, presence of SiC provides good surface passivation. The Ray tracing and Finite-difference time-domain method is used for the computation of the optoelectronics properties. Also, total current density from *P*+–*I*–*N*+ region is presented in Fig. 7. Higher current density is observed in case of FDTD simulation compared to ray tracing method. This is because the ray tracing algorithm does not includes the diffraction and coherence effects, whereas FDTD algorithm accounts both diffraction and coherence effect. This algorithm solves the electric and magnetic field propagation directly from the fundamental wave equations. FDTD simulation results in a higher carrier generation inside the device. The light generated current is directly proportional to the light generated carriers.

3.4 Electrical characteristic i.e., J-V characteristics of $SiC/Si_3N_4/SiO_2$ passivated BC-BJ SiGe solar cell

The most important part of the solar cell design is its current voltage characteristics. Any design must have a higher short circuit current density (J_{sc}) and a higher open circuit voltage (V_{oc}). The current density–voltage (J-V) characteristic of the device is measured under AM1.5 illumination with optical power (0.1 W/cm^2) and is presented in Fig. 8 along with power density- voltage (P_o-V) curve. Ray tracing and FDTD simulation has been done to obtained the J-V curve. FDTD simulation has been done in order to obtain the maximum possible efficiency from the device. In case of FDTD, zero optical lose is considered. This shows, in future, further enhancement in absorption of device will lead to ~18 % PCE. In FDTD, We obtained the short circuit current density (J_{sc}) of 37.86 mA·cm^{-2} and open circuit voltage (V_{oc}) of 593 mV, whereas in ray tracing J_{sc} and V_{oc} are 28.93 mA·cm^{-2} and 585 mV, respectively. FDTD

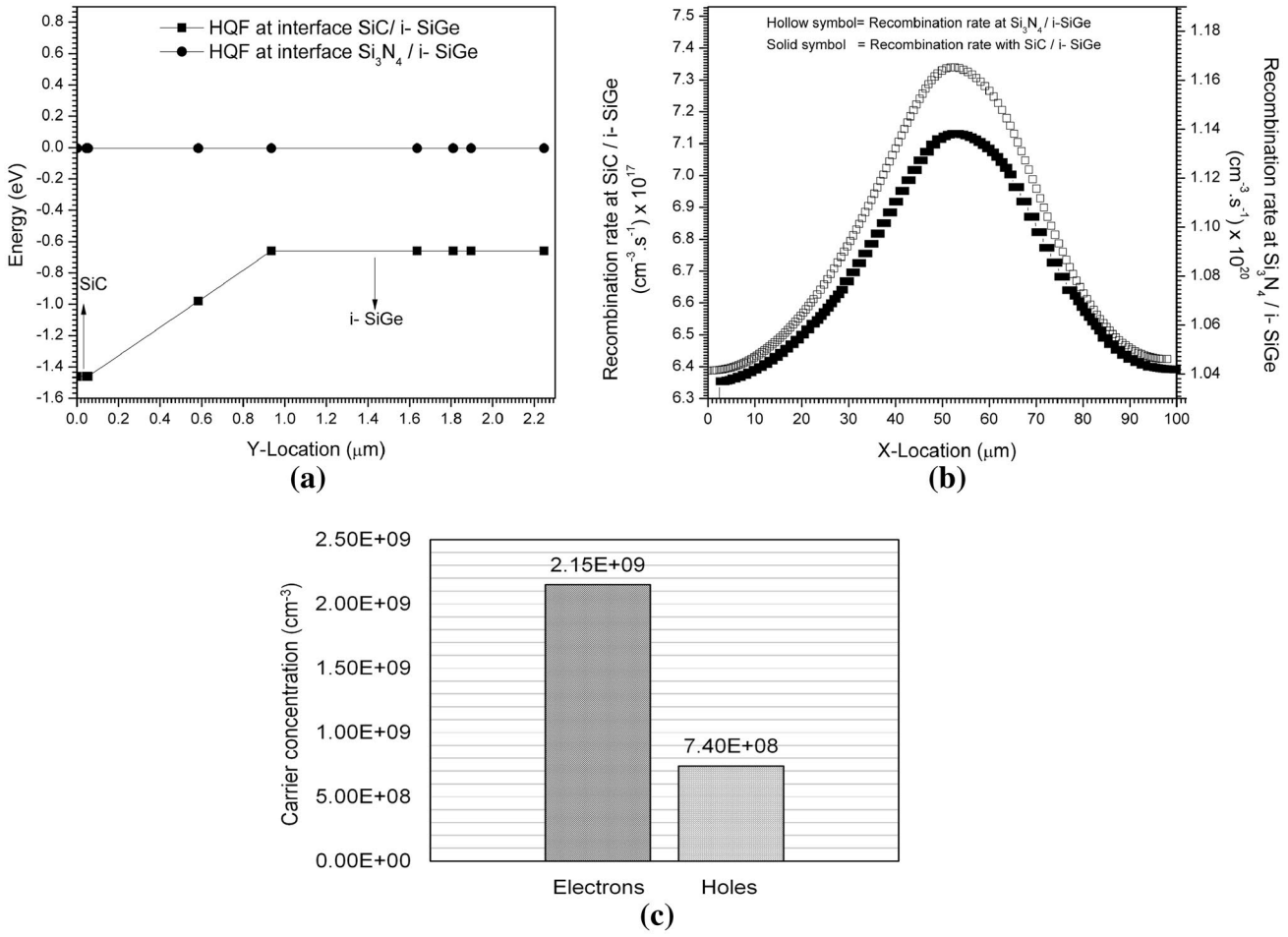


Fig. 6 Advantageous parameters at the interface. **a** Comparison of hole quasi Fermi (HQF) energy levels. **b** Recombination rate near *SiC/I-SiGe* interface, primary axis and recombination rate near *Si_3N_4/I-SiGe* interface, secondary axis. **c** Electrons and holes concentrations near *SiC/I-SiGe* interface

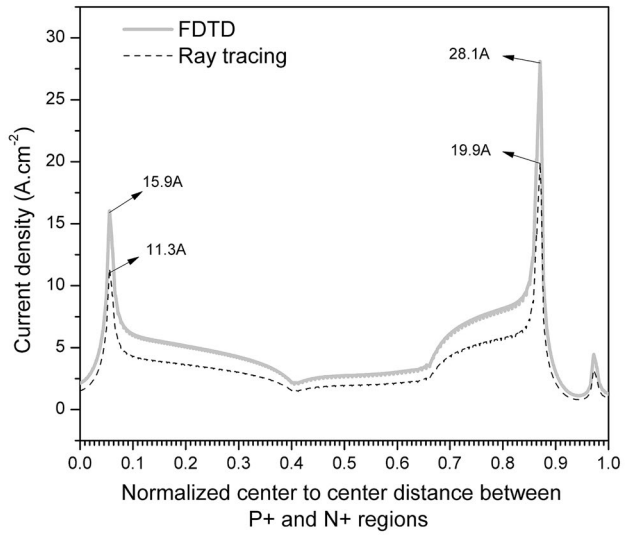


Fig. 7 Comparison of light generated current density of BC-BJ SiGe solar cell using ray tracing and FDTD method

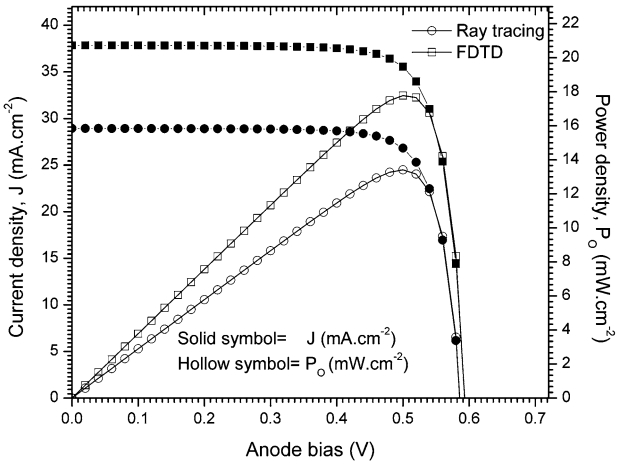


Fig. 8 Electrical parameters of $SiC/Si_3N_4/SiO_2$ passivated BC-BJ SiGe Solar Cell: J-V curve-primary axis and P_o-V curve-secondary axis

Table 2 Photovoltaic properties of ultra-thin BC-BJ SiGe Solar cell

	J_{sc} (mA·cm^{-2})	V_{oc} (V)	FF (%)	PCE (%)
FDTD	37.8	0.593	79.1	17.8
Ray tracing	28.9	0.585	79.1	14.8

FF fill factor, *PCE* power conversion efficiency

simulation shows 30.8 and 1.4 % increment in J_{sc} and V_{oc}, respectively compared to ray tracing simulation. The FDTD simulation results in higher absorption and higher carrier concentration which further results in higher Jsc and variation in the V_{oc} is negligible. Some of the important parameters of the design have been summarized in Table 2.

4 Conclusion

In this work, we proposed a novel, thin BC-BJ SiGe Solar cell design as a cost-effective solution for energy efficient applications. Results, indicates that presence of SiC play an important role in photoelectric conversion, photo reflection is significantly reduced in UV/Visible spectrum region in case of SiC/Si_3N_4/SiO_2 passivated BC-BJ SiGe solar cell. The work analyses the band diagrams, internal current density, and surface recombination of the proposed design. The ray tracing and FDTD method is used to simulate the device optoelectronics characteristics. TCAD analysis results in a higher EQE >82 % throughout the spectrum range of 450–700 nm wavelength and a significantly improved power conversion efficiency of 14.8 %. The carrier concentration at the interface is lower than the bulk, therefore reduced recombination rate of ~10^{17} cm^{-3} s^{-1} has been obtained at the interface. We have achieved the concentration of ~10^{9} cm^{-3} and ~10^{8} cm^{-3} for electron and hole, respectively at the interface. This ensures the enhanced surface passivation. In addition, enhanced photovoltaic properties: J_{sc} (28.9 mA), V_{oc} (0.585 V), FF (79.1 %) are also observed for the proposed design. Improved photovoltaic properties indicates proposed device with SiC passivation can be used in lightweight, low cost installation of PV modules. Also, device has rear contact architecture it is easier to interconnect and can be placed closer together in the modules, no need of space between cells. Therefore, it can support lightweight, low cost installation of PV modules. Also, rear contact architecture supports its application in concentrator photovoltaics (CPV) system. Result concludes, proposed device can be a cost effective candidate for PV applications.

Acknowledgments The authors would like to thank Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University to carry out this work. Rahul Pandey (JRF) acknowledges UGC, Govt. of India for providing fellowship.

References

Aberle AG (2010) Surface passivation of crystalline silicon solar cell a review. Prog Photovolt Res Appl 8(5):487–473. http://onlinelibrary.wiley.com/doi/10.1002/1099-159X(200009/10)8:5%3C473::AID-PIP337%3E3.0.CO;2-D/pdf. Accessed 14 May 2015

Franklin E et al (2014) Design, fabrication and characterisation of a 24.4 % efficient interdigitated back contact solar cell. Prog Photovolt Res Appl. doi:10.1002/pip.2556

Atlas user manual (2014) Silvaco, Inc. 4701 Patrick Henry Drive, Bldg. Santa Clara, CA 95054

Cruz-Campa et al (2011) Microsystems enabled photovoltaics: 14.9 % efficient 14 μm thick crystalline silicon solar cell. Sol Energy Mater Sol Cells 95(2):551–558. doi:10.1016/j.solmat.2010.09.015

Deceglie MG, Ferry VE, Alivisatos AP, Atwater HA (2012) Design of nanostructured solar cells using coupled optical and electrical modeling. Nano Lett 12(6):2894–2900. doi:10.1021/nl300483y

Deinega A, Eyderman S, John S (2013) Coupled optical and electrical modeling of solar cell based on conical pore silicon photonic crystals. J Appl Phys 113:224501

Dziewior J, Schmid W (1977) Auger coefficient for highly doped and highly excited silicon. Appl Phys Lett 31(5):346–348

Feldmann F, Bivour M, Reichel C, Hermle M, Glunz SW (2014) Passivated rear contacts for high-efficiency n-type Si solar cells providing high interface passivation quality and excellent transport characteristics. Sol Energy Mater Solar Cells 120:270–274. doi:10.1016/j.solmat.2013.09.017

Garnett E, Yang P (2010) Light trapping in silicon nanowire solar cells. Nano Lett 10:1082–1108. doi:10.1021/nl100161z

Green MA, Emery K, Hishikawa Y, Warta W, Dunlop ED (2014) Solar cell efficiency tables 43. Prog Photovolt Res Appl. doi:10.1002/pip.2452

Hall RN (1952) Electron-hole recombination in germanium. Phys Rev 87(2):387. doi:10.1103/PhysRev.87.387

Jeong S, McGehee MD, Cui Y (2013) All-back-contact ultra-thin silicon nanocone solar cells with 13.7 % power conversion efficiency. Nature Communications 4. Article number 2950:1–7. doi:10.1038/ncomms3950

Kerschaver EV, Beaucarne G (2006) Back-contact solar cells: a review. Prog Photovolts 14(2):107–123. doi:10.1002/pip.657

Ouellete J (2002) Silicon-Germanium Gives Semiconductor the Edge. Am Inst Phys pp 22–25. https://prezi.com/odedrjji90jc/will-hambleton-silicon-germanium-gives-semiconductors-the-edge/

Pandey R, Chaujar R (2014) Enhanced Back-Contact Back-Junction Crystalline Silicon Solar Cell Performance With A Silicon-Carbide (SiC) Based Front Surface Passivation. Int J Adv Technol Eng Sci 2(01):626–630. http://www.ijates.com/images/short_pdf/1412233437_151.pdf. Accessed 14 May 2015

Pandey R, Chaujar R (2014) Front Surface Passivation Scheme for Back-Contact Back-Junction (BC-BJ) Silicon Solar Cell. International Conference on Nanotechnology, Nanocon 014, Pune, India

Povolny H, Agarwal P, Han S, Deng X (2000) Comparison study of a-SiGe solar cells and materials deposited Using different hydrogen dilution. Mat Res Soc Symp Proc 609:6. doi:10.1557/PROC-609-A30.3

Schwartz RJ, Lammert MD (1975) Silicon solar cells for high concentration applications. IEEE International Electron Devices Meeting, pp 350–352. doi:10.1109/IEDM.1975.188896

Selberherr S (1984) Analysis and simulation of semiconductor devices. Springer, Vienna, New York. doi:10.1007/978-3-7091-8752-4

Shockley W, Read WT (1952) Statistics of the recombinations of holes and electrons. Phys Rev 87(5):835–842. doi:10.1103/PhysRev.87.835

Slotboom JW (1977) The pn-product in silicon. Solid State Electron 20(4):279–283. doi:10.1016/0038-1101(77)90108-3

Slotboom JW, de Graaff HC (1976) Measurements of bandgap narrowing in Si bipolar transistors. Solid State Electron 19(10):857–862. doi:10.1016/0038-1101(76)90043-5

Tian B et al (2007) Coaxial silicon nanowires as solar cells and nanoelectronic power sources. Nature 449:885–889. doi:10.1038/nature06181

Tomasi A et al (2014) Back-Contacted Silicon Heterojunction Solar Cells With Efficiency >21%. IEEE J Photovolt 4(4):1046–1054. doi:10.1109/JPHOTOV.2014.2320586

Wang CC et al (2012) Characterization of nanocrystalline SiGe thin film solar cell with double graded-dead absorption layer. Int J Photoenergy. doi:10.1155/2012/890284

Yang D, Yu X, Li X, Wang P, Wang L (2010) Germanium-doped crystal silicon for solar cells. Solid-State and Integrated Circuit Technology (ICSICT), 10th IEEE International Conference. doi:10.1109/ICSICT.2010.5667855

On the Assessment of Segmentation Methods for Images of Mosaics

Gianfranco Fenu[1], Nikita Jain[2], Eric Medvet[1],
Felice Andrea Pellegrino[1] and Myriam Pilutti Namer[3]
[1]*Department of Engineering and Architecture, University of Trieste, Italy*
[2]*Department of Software Engineering, Delhi Technological University, India*
[3]*Department of Humanities, Ca' Foscari University of Venice, Italy*
{*fenu, emedvet, fapellegrino*}*@units.it*

Keywords:
Image segmentation, superpixel, comparative evaluation, cultural heritage preservation.

Abstract:
The present paper deals with automatic segmentation of mosaics, whose aim is obtaining a digital representation of the mosaic where the shape of each tile is recovered. This is an important step, for instance, for preserving ancient mosaics. By using a ground-truth consisting of a set of manually annotated mosaics, we objectively compare the performance of some existing recent segmentation methods, based on a simple error metric taking into account precision, recall and the error on the number of tiles. Moreover, we introduce some mosaic-specific hardness estimators (namely some indexes of how difficult is the task of segmenting a particular mosaic image). The results show that the only segmentation algorithm specifically designed for mosaics performs better than the general purpose algorithms. However, the problem of segmentation of mosaics appears still partially unresolved and further work is needed for exploiting the specificity of mosaics in designing new segmentation algorithms.

1 Introduction and related work

Mosaics are artworks reproducing decorative images or patterns out of small colored tiles (called *tesserae* or *tessellae*). Ancient mosaics represent a relevant part of the cultural heritage of many countries and digital imaging may provide tools for their preservation, distant consultation, reconstruction, anastylosis (Benyoussef and Derrode, 2011). Traditionally, for preservation purposes, ancient mosaics are manually acquired by tracing the contour of each tile over a semi-transparent paper superimposed to the mosaic. Needless to say, this is a time-consuming task but due to the specificity of mosaic, it is important to get the shape of each tile. Another possibility is acquiring a digital image of the mosaic and applying an image segmentation algorithm to automatically detect the tiles thus building a "digital model" of the mosaic.

Dispose of a digital model of the mosaic could be a great advantage both in the traditional fields referring to the humanities and in the new technologies applied to the study, the preservation and the valorization of the worldwide cultural heritage. Indeed, a "digital model" could be useful for archaeologists, scholars and restorers who are interested in studying, comparing and preserving the mosaics. Moreover, for restorers a digital model of a mosaic could become an essential professional tool. Coming to the new technologies applied to the cultural heritage, the virtual reproduction of a mosaic is the first step to build the 3D model. This could be used to rebuild the place where the mosaic or the mosaics was/were originally set, helping in visualizing it for preservation and valorization use.

To the best of our knowledge, a single mosaic-oriented segmentation algorithm has been proposed in the literature (Benyoussef and Derrode, 2008) that is based on the well-known watershed algorithm (Vincent and Soille, 1991) and some mosaic-specific preprocessing. Unfortunately, its performance has been assessed subjectively only.

On the other hand a number of recent general purpose segmentation algorithms exist (Moore et al., 2008; Levinshtein et al., 2009; Liu et al., 2011; Achanta et al., 2012), but still is not clear whether they are suitable for segmenting mosaics and how they comparatively perform. The purpose of the present paper is twofold: on the one hand, we suggest that a simple error metric taking into account precision, recall and the error on the number of tiles is appropriate for the objective evaluation of a mosaic segmentation. We further introduce some mosaic-specific hardness estimators (namely some indexes of how difficult is the task of segmenting a particular mosaic image). On the other hand, we compare the performance of a number of popular segmentation algorithms based on a data-set that we collected and manually annotated thus providing a ground-truth that could be employed to assess performance of new mosaic segmentation algorithms.

The history of image segmentation evaluation dates back to the late seventies (Yasnoff et al., 1977). From then on, a plethora of evaluation methods, of both supervised and unsupervised nature have been proposed, see (Zhang et al., 2008) for a relatively recent survey. Here we employ a supervised (i.e., based on manually segmented images) evaluation method similar to the precision and recall method applied by (Neubert and Protzel, 2012), but with reference to the whole segment area instead of its boundary. Moreover, since the number of segments is an objective datum of the problem, it is reasonable to take into account the difference of the number of detected and ground-truth segments.

A survey, even partial, of segmentation methods is far beyond the purpose of this introduction. Here we only stress that a single segmentation method (Benyoussef and Derrode, 2008) specifically designed for mosaics has been proposed in the literature. The results reported in Section 3.1 have been obtained by using the original code kindly provided by the authors. As far as the other image segmentation algorithms employed in this paper are concerned, they all belong to the class of superpixel segmentation methods (Ren and Malik, 2003) meaning that the segmentation criterion is perceptual uniformity. In particular, we employed the Superpixel Lattice (SL) (Moore et al., 2008), the TurboPixels (TP) (Levinshtein et al., 2009), the Entropy Rate Superpixel Segmentation (ERSS) (Liu et al., 2011) and the SLIC Superpixels (SLIC) (Achanta et al., 2012). The segmentation methods will be briefly described in Section 3.1.

The reminder of the paper is organized as follows: in Section 2 the performance metrics and the hardness estimators are introduced. In Section 3.1 the segmentation methods, the dataset and the experimental results are reported. Finally, the conclusions are drawn in Section 4.

2 Objective evaluation metrics

Image segmentation is the process of partitioning a digital image into segments, representing possibly meaningful regions of the image. For the purpose of the present paper, a *segmentation method* $\mathcal{S}$ is a function such that $\mathcal{S}(p)$ is the label assigned to pixel $p \in I$, for a given image I. We allow a special case when the method deems the pixel not to belong to any meaningful region: we denote this case with $\mathcal{S}(p) = \varnothing$. Indeed, we believe that a segmentation method which is tailored to mosaic images should assign a $\varnothing$ label to the pixels corresponding to the filler, if any. A *region* $R \subseteq I$ is the largest set of pixels of a segmented image for which the label is the same and different from $\varnothing$. The *segmentation* $\mathcal{R}_I$ of the image I is the set of all regions.

A *ground truth* $\mathcal{T}_I$ is a set of non-overlapping pixel subsets of an image I; $\mathcal{T}_I$ represents the desired segmentation of a mosaic image. A *tile* $T \in \mathcal{T}_I$, with $T \subseteq I$, contains all and only the pixels corresponding to the same tessella on the mosaic depicted in I.

2.1 Performance metrics

Given a segmentation $\mathcal{R}_I$ performed by a method $\mathcal{S}$ on an image I and the desired segmentation $\mathcal{T}_I$, we aim at determining quantitatively how close $\mathcal{R}_I$ is to $\mathcal{T}_I$. To this end, we define three metrics: average tile precision $\mathrm{Prec}(\mathcal{R}_I, \mathcal{T}_I)$, average tile recall $\mathrm{Rec}(\mathcal{R}_I, \mathcal{T}_I)$, tile count error $\mathrm{Count}(\mathcal{R}_I, \mathcal{T}_I)$. More in details:

$$\mathrm{Prec}(\mathcal{R}_I, \mathcal{T}_I) = \frac{1}{|\mathcal{T}_I|} \sum_{T \in \mathcal{T}_I} \frac{\max_{R \in \mathcal{R}_I} |R \cap T|}{|R|}$$

$$\mathrm{Rec}(\mathcal{R}_I, \mathcal{T}_I) = \frac{1}{|\mathcal{T}_I|} \sum_{T \in \mathcal{T}_I} \frac{\max_{R \in \mathcal{R}_I} |R \cap T|}{|T|}$$

$$\mathrm{Count}(\mathcal{R}_I, \mathcal{T}_I) = \frac{\mathrm{abs}(|\mathcal{T}_I| - |\mathcal{R}_I|)}{|\mathcal{T}_I|}$$

We chose to use precision and recall because the problem of segmenting mosaic images can

be seen, to some degree, as an information retrieval problem: that is, given a mosaic tessella depicted on the image, how good is the segmentation method in finding all (recall) and only (precision) the relevant pixels, i.e., those which correspond to that tessella? In other words, precision and recall assess the method ability to correctly cover a tile with a region.

However, precision and recall alone cannot completely capture the segmentation method ability to meet the objective described in Section 1. For example, consider the case in which each tile in $\mathcal{T}_I$ is perfectly coincident to a region in $\mathcal{R}_I$, but there are several spurious regions in $\mathcal{R}_I$ which do not cover any tile—i.e., they cover the filler. Clearly, these regions will negatively affect the result of simple analysis of the mosaic performed on the segmented image, such as counting the number of tessellas or determining the amount of filler. Hence, in order to tackle this limitation, we defined the tile count error.

Finally, we also use the F-measure index, which is the harmonic mean of precision and recall:

$$\mathrm{Fm}(\mathcal{R}_I, \mathcal{T}_I) = 2\frac{\mathrm{Prec}(\mathcal{R}_I, \mathcal{T}_I)\mathrm{Rec}(\mathcal{R}_I, \mathcal{T}_I)}{\mathrm{Prec}(\mathcal{R}_I, \mathcal{T}_I) + \mathrm{Rec}(\mathcal{R}_I, \mathcal{T}_I)}$$

2.2 Hardness estimators

We aim at quantifying those style-related respects characterizing a mosaic image which likely affect how difficult is the task of segmenting a particular mosaic image, i.e., at defining some *hardness estimators*. We considered four style features—tiles size variability, amount of filler, color dissimilarity within the tiles, color dissimilarity near to the edge of the tiles—and the corresponding indexes, which are defined in details below. We think that, in general, i.e., without making any assumption about the segmentation method which is being used, the lower the index, the easier the image segmentation—except for the last, for which the greater the index, the easier the segmentation.

We chose to consider the amount of filler as an hardness estimator because most of existing image segmentation methods usually do not allow for an image portion to not be covered by any region: however, when operating on mosaic images, this should be the ideal behavior for those image portions which corresponding to the filler. Hence, we think that the amount of filler may significantly impact on segmentation.

Of course, other image features may impact on segmentation hardness: e.g., sharpness, geometric distortion, resolution, etc. In this work, we focus on mosaic-related characteristics.

Each index can be computed starting from a segmentation of the image: in the following notation, we assume that the indexes are computed using the ground truth $\mathcal{T}$. We are not concerned here in providing some proxy for these indexes which can be computed without a segmentation.

Tile size variability We define the index $\rho_{|T|} = \frac{\sigma_{|T|}}{\mu_{|T|}}$, where $\mu_{|T|}$ and $\sigma_{|T|}$ are the mean and standard deviations of the tiles size $|T|$, respectively, with $T \in \mathcal{T}$. $\rho_{|T|}$ captures the variability of tile size.

Filler amount We define the index $\rho_{\varnothing} = 1 - \frac{\sum_{T\in\mathcal{T}|T|}}{|I|}$, i.e., the relative size of the portion of the image which is not covered by any tile.

In-tile color dissimilarity We define the index μ_{I_T} as the mean of the in-tile color dissimilarity $I_T = \sigma_{L^*} + \sigma_{a^*} + \sigma_{b^*}$, where σ_{L^*}, σ_{a^*} and σ_{b^*} are the standard deviation of the L^*, a^* and b^* channel values of the pixels within the tile T (in the CIE-Lab color space). I_T captures the dissimilarity of the color within a tile T and could be affected (i.e., have a large value) by deterioration, tile surface unevenness causing shadows, and so on.

Out-tile color dissimilarity We define the index μ_{O_T} as the mean of the out-tile color dissimilarity O_T, which is computed, for each tile T, as follows. Let $\bar{T}$ be a supertile of T which has the same barycenter of T and is scaled by a factor $\beta > 1$. We set:

$$O_T = \Bigg(\left(\mu_{L^*}^{T} - \mu_{L^*}^{\bar{T}\setminus T}\right)^2 + \left(\mu_{a^*}^{T} - \mu_{a^*}^{\bar{T}\setminus T}\right)^2 + + \left(\mu_{b^*}^{T} - \mu_{b^*}^{\bar{T}\setminus T}\right)^2 \Bigg)^{\frac{1}{2}}$$

where $\mu_{L^*}^{T}$ and $\mu_{L^*}^{\bar{T}\setminus T}$ are the mean values of the L^* channel values in the pixel of T and $\bar{T} \setminus T$, respectively, being $\bar{T} \setminus T$ the portion of the supertile $\bar{T}$ not covered by T (the same for a^* and b^*). O_T captures the dissimilarity of the mean color of the tile T from the mean color of the area adjacent to T. We set $\beta = 1.2$.

3 Experimental comparison

3.1 Segmentation methods

We considered 5 segmentation methods. One of them (Benyoussef and Derrode, 2008) is tailored specifically to mosaics and will be referred to as Tessella-Oriented Segmentation (TOS). The remaining 4 methods are of superpixel type, namely they aim to produce segments that are perceptually uniform. Due to the nature of mosaics, this kind of segmentation is desirable. The methods have been selected among:

- recent (posterior to 2008);
- fast (meaning that, with reasonable parameters and reasonable image size, the processing time is no more than a few minutes—for instance, the processing time[1] of SLIC on a 647 kpixel image is 83 s);
- whose code is publicly available for free.

Based on the above requirements we selected the Superpixel Lattice (SL) (Moore et al., 2008), the TurboPixels (TP) (Levinshtein et al., 2009), the Entropy Rate Superpixel Segmentation (ERSS) (Liu et al., 2011) and the SLIC Superpixels (SLIC) (Achanta et al., 2012).

3.1.1 Tessella-Oriented Segmentation (TOS)

The main idea of the algorithm is modelling explicitly the mortar joint, i.e., exploiting the fact that the tesserae are surrounded by the filler. More precisely, for each pixel, some directional gray-level profiles are calculated and compared to a parametrized model. In this way a "criterion image" to be processed by the watershed transform is obtained. The algorithm has a single parameter (α) corresponding to half of the average tile size.

3.1.2 Superpixel Lattice (SL)

The algorithm relies on subsequent bipartitions of the image. The bipartitions are, alternatively, horizontal and vertical and such that the topology of a regular lattice is preserved. The optimal partition paths are chosen in order to maximize the cross-part difference, hence are likely to pass through actual image boundaries. The main parameter is the *resolution*, corresponding to the total number of superpixels, organized in a $l_w \times l_h$ grid.

3.1.3 TurboPixels (TP)

TP is a geometric flow-based algorithm, where K circular seeds are distributed uniformly over the image and their boundaries evolve in time according to a diffusion-like equation. The evolution is governed by conflicting terms, called reaction-diffusion and "doublet", whose weights α and β respectively, have been assigned the values suggested by the authors. Thus, the algorithm depends on the parameter K only.

3.1.4 Entropy Rate Superpixel Segmentation (ERSS)

In the ERSS algorithm, the whole image is represented by a graph, whose nodes correspond to the pixels, while the arc weights between adjacent nodes denote the pairwise similarities. Then, the segmentation problem is formulated as a graph partition problem, where K subgraphs (corresponding to K superpixels) are found in order to maximize a proper objective function. The objective function comprises two terms: the entropy rate of a random walk on a graph and a balancing term. The entropy rate favors formation homogeneous clusters, while the balancing function enforces clusters of similar size. Besides K the algorithm relies on two parameters: the weight λ' of the balancing term w.r.t. the entropy rate term and the width σ of the Gaussian kernel used for computing pixel similarity. According to the authors, $\lambda' = 0.5$ yields a good compromise between the opposite requirements, while the algorithm is little sensitive to σ for a wide range of values. Hence the most significant parameter is K.

3.1.5 SLIC Superpixels (SLIC)

SLIC stands for Simple Linear Iterative Clustering and is a simple adaptation of the well-known k-means clustering algorithm. The distance measure employed combines color and spatial proximity. In order to speed-up calculations, the distances are computed in a neighborhood proportional to the superpixel size. The algorithm has two parameters, precisely the desired number of approximately equally sized superpixels and the weighting factor between colour and spatial proximity for the calculation of the distance measure.

[1] Mean execution time of the SLIC segmentation algorithm, using Matlab R2012a on MacBook Pro, with 8 GB RAM and CPU Intel Core 2 Duo, 3.06 GHz.

Based on extensive experimentation, the main parameter has been identified as the weighting factor w. Indeed, we noted that different weighting factor values may result in segmentations with very different number of superpixels, even if using the same initial guess for the desired number of superpixels.

3.2 Mosaics

We considered 5 mosaic images, depicting 5 mosaics which differ in age and style (see Figure 1):

Church This is a portion of a mosaic forming the floor of the Basilica di Santa Maria Assunta, the principal church in the town of Aquileia (Udine, Italy). The original church dated back to the fourth century; the current basilica was built in the eleventh century and rebuilt again in the thirteenth century. The image has been acquired with a camera not perfectly orthogonal to the mosaic plane.

Museum This is a portion of a mosaic which is displayed at the Early Christian Museum (Museo paleocristiano di Monastero) in Aquileia. The museum houses several mosaic fragments from different excavations of Aquileia.

Bird This is a contemporary mosaic built by an Italian artist[2] inspired by mosaics of the aforementioned Basilica.

Flower This is a small contemporary mosaic, built by an Italian amateur as an essay for a course of ancient mosaic technique.

University This is a portion of a mosaic forming the floor in the main building of our campus, which dates back to the 1938. This mosaic is, in several points, quite deteriorated due to pounding.

Table 1 shows the values of the hardness estimators indexes (leftmost 4 columns) for the 5 mosaic images, along with width and height of the image, the number $|\mathcal{T}_I|$ of tiles in the corresponding ground truth, the size of the image $|I|$, and the average size of a tile $\mu_{|T|}$.

3.3 Results and discussion

We applied each method to each mosaic image, with several values for the main method parameter (totaling 603 method applications) and obtained the corresponding segmentations: with respect to the main parameter value, we sampled more densely those intervals in which the performance appeared to be better. We then computed the performance metrics presented in Section 2.

With the aim to harmonize the results and to easily compare the performances of different algorithm, we chose to scale the main parameter of each method (as briefly identified and described above) by the effective number of tiles in each mosaic, when appropriate. Indeed, the parameters of TP, ERSS and SL methods are the desired numbers of superpixels, hence scaling the parameters allows for performance comparison of every method applied on the full set of available mosaics. Instead, considering the other two algorithms (TOS and SLIC), the parameters are related to different mosaic information and the scaling by the effective tiles number would be useless. More in detail, let p be the main parameter: we set $p := \alpha$ for TOS, $p := \frac{l_w l_h}{|\mathcal{T}_I|}$ for SL, $p := \frac{K}{|\mathcal{T}_I|}$ for TP and ERSS, and $p := w$ for SLIC. For SL, we experimented, for each mosaic image, with l_w, l_h values consistent with the width/height ratio of the image.

The results are shown graphically in Figure 2 and are summarized in Table 2. Figure 2 consists of 20 charts, one row per method: the first column (left) shows the F-measure vs. the main parameter, the second column (center) shows the count error vs. the main parameter, and the third column (right) shows F-measure vs. count error. Charts of the first two columns show one line per mosaic image. Table 2 shows the performance metrics obtained by the methods on the mosaic images, using only the main parameter value which led to the best F-measure.

It can be seen that the best (largest) values in terms of F-measure are obtained with TOS and SLIC methods. In particular, these methods achieve 74.1% and 77.1% F-measure for the Museum image—best values for ERSS, TP and SL are respectively 46.3%, 67.9% and 66.6% on the Bird, University and Museum images. This finding is confirmed by Table 2, which highlights, for each mosaic, the best values obtained for the count error and the F-measure: concerning the latter, SLIC performs best on 2 images, TP on 2 and TOS on the last image; TOS is the second best performing method in 3 images.

SL and TOS are the best performing methods in terms of the count error—note also the range of the y-axis of the corresponding charts. We recall that, for the former, the main parameter deter-

[2] `http://latenagliaimpazzita.wordpress.com/2014/03/04/aquileia-eia-eia-uh`



(a) Church (b) Museum (c) Bird (d) Flower (e) University

Figure 1: The mosaics images considered in our evaluation.

Mosaic image	Size	$\lvert\mathcal{T}_I\rvert$	$\lvert I\rvert$	$\mu_{\lvert T\rvert}$	$\rho_{\lvert T\rvert}$ [%]	$\rho_{\varnothing}$ [%]	μ_{I_T}	μ_{O_T}
Church	793×1031	855	0.82×10^6	543.5	59.5	43.2	39.2	806.9
Museum	1951×2386	135	4.66×10^6	21 451.8	40.1	37.8	13.8	201.0
Bird	942×704	717	0.66×10^6	530.0	58.0	42.7	54.3	618.2
Flower	1097×1872	546	2.05×10^6	2646.0	41.8	29.6	42.6	315.5
University	438×340	134	0.15×10^6	666.1	70.3	40.1	35.1	962.3

Table 1: Salient information of the 5 considered mosaic images.

Mosaic image	Method	p	Count	Prec	Rec	Fm
Church	TOS	4.00	0.54	56.4	71.9	63.2
	SL	1.08	**0.08**	35.3	71.1	47.2
	TP	2.39	1.54	63.5	63.8	**63.7**
	ERSS	2.83	1.83	51.2	39.3	44.4
	SLIC	12.83	0.25	47.5	83.8	60.6
Museum	TOS	6.00	0.14	64.4	87.3	74.1
	SL	1.07	**0.07**	52.3	91.7	66.6
	TP	1.73	0.84	60.8	60.3	60.6
	ERSS	2.16	1.18	48.0	39.5	43.3
	SLIC	5.00	7.03	73.9	80.6	**77.1**
Bird	TOS	4.00	**0.03**	52.8	81.7	64.2
	SL	1.31	0.31	40.4	68.6	50.9
	TP	2.14	1.28	62.9	68.4	**65.5**
	ERSS	2.94	1.94	55.3	39.9	46.3
	SLIC	12.00	0.32	50.6	84.8	63.4
Flower	TOS	4.00	0.06	49.4	67.8	57.2
	SL	0.99	**0.01**	37.1	63.1	46.7
	TP	1.72	0.83	67.4	63.5	65.4
	ERSS	2.75	1.75	60.2	35.3	44.5
	SLIC	6.00	1.74	63.4	70.4	**66.8**
University	TOS	4.00	0.90	63.2	78.5	**70.1**
	SL	1.04	**0.04**	44.0	88.6	58.8
	TP	1.90	1.78	67.0	68.9	67.9
	ERSS	2.99	1.99	52.6	40.9	46.0
	SLIC	8.00	0.33	54.5	87.9	67.3

Table 2: Performance metrics of the segmentation methods obtained on the mosaic images using the main parameter value (third column) which led to the best F-measure. For each mosaic image, the best value for the count error and the best value for the F-measure are highlighted.

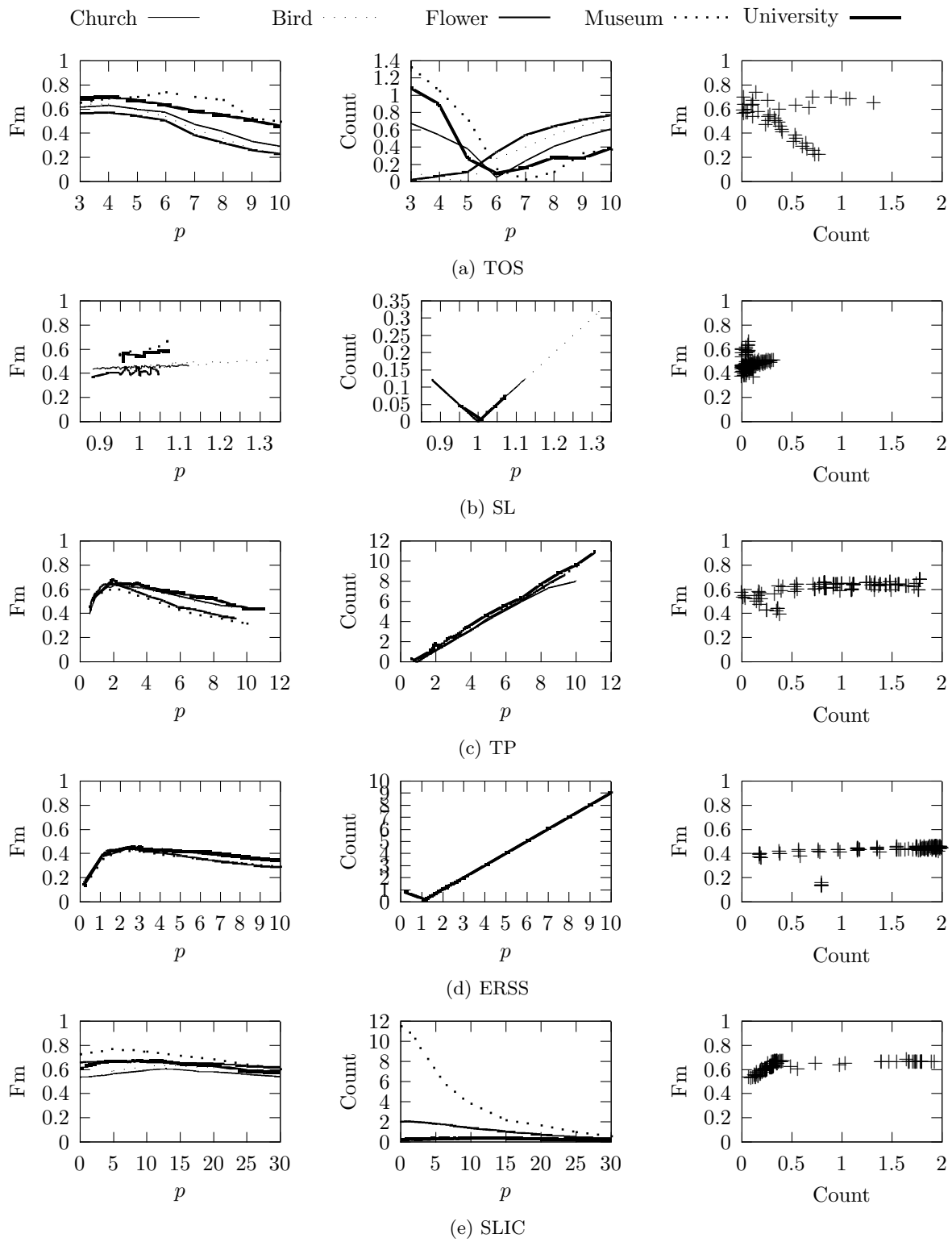


Figure 2: F-measure (left column), count error (central column) for the 5 methods (one row per method) on the 5 mosaic images (one line per image) vs. the methods main parameter (see text). The third column shows F-measure vs. count error: top-left corner is the ideal performance.

mines exactly which will be the number of regions in the segmentation: hence, setting $p = \frac{l_w l_h}{|\mathcal{T}_I|} = 1$ leads to a count error equal to 0. However, in a real scenario $|\mathcal{T}_I|$ could not be available, or could be estimated only roughly. On the other hand, for the TOS method, the main parameter does not depend on $|\mathcal{T}_I|$; morevoer the best values for the F-measure and count error metrics are obtained for (approximately) the same values for the main parameter.

This finding is further corroborated by the Fm vs. Count chart (third column of Figure 2): for the TOS method, the points with the largest F-measure and those with the smallest count error are quite close. Instead, for the other methods (in particular for the SLIC method), this chart shows a clear trade-off between F-measure and count error: in other words, the main parameter value which allows to obtain a good segmentation in terms of F-measure, also leads to a remarkably bad result in terms of count error (50% more regions found than the actual tiles). We think that this limitation is caused by the fact that methods other than TOS do not take into account the presence of the filler. Hence, they tend to introduce regions which cover the filler but negatively affect the count error. An example of this effect can be viewed in Figure 3, which shows the best segmentations (in terms of F-measure) obtained by the four methods on the University mosaic image, along with (Figure 3(a)) the ground truth.

Concerning the sensitivity of the methods w.r.t. their main parameters, results show that the count error is affected more distinctly than the F-measure. We remark that for 3 on 5 methods (SL, TP and ERSS) the main parameter represents the desired number of regions in the segmentation, i.e., the number of tiles: a reliable estimate for this figure could not be available in a real world scenario. For the other two methods (TOS and SLIC), performances apper to be little sensitive to reasonable variations in the main parameter.

Another interesting result concerns how the hardness estimators (see Section 2.2) impact on the method performances. The two best performing methods (TOS and SLIC) obtain the best F-measure values on the Museum and University images. Indeed, the Museaum image has the smallest value for the μ_{I_T} index and the University image has the largest value for the μ_{O_T} index (see Table 1). Despite the number of mosaic images which have been analyzed is not sufficient for drawing statistically significant conclusions, this finding appears to support the hypothesis that μ_{I_T} and μ_{O_T} can be used as good predictors of the performance of a mosaic image segmentation method. From another point of view, if one would aim at improving the performance of a segmentation method, he/she could augment the method with an image preprocessing step which decreases the value of μ_{I_T} and increases the value of μ_{O_T}. Recall, however, that the actual values for μ_{I_T} and μ_{O_T} cannot be computed without knowing the ground truth of the mosaic image to be analyzed.

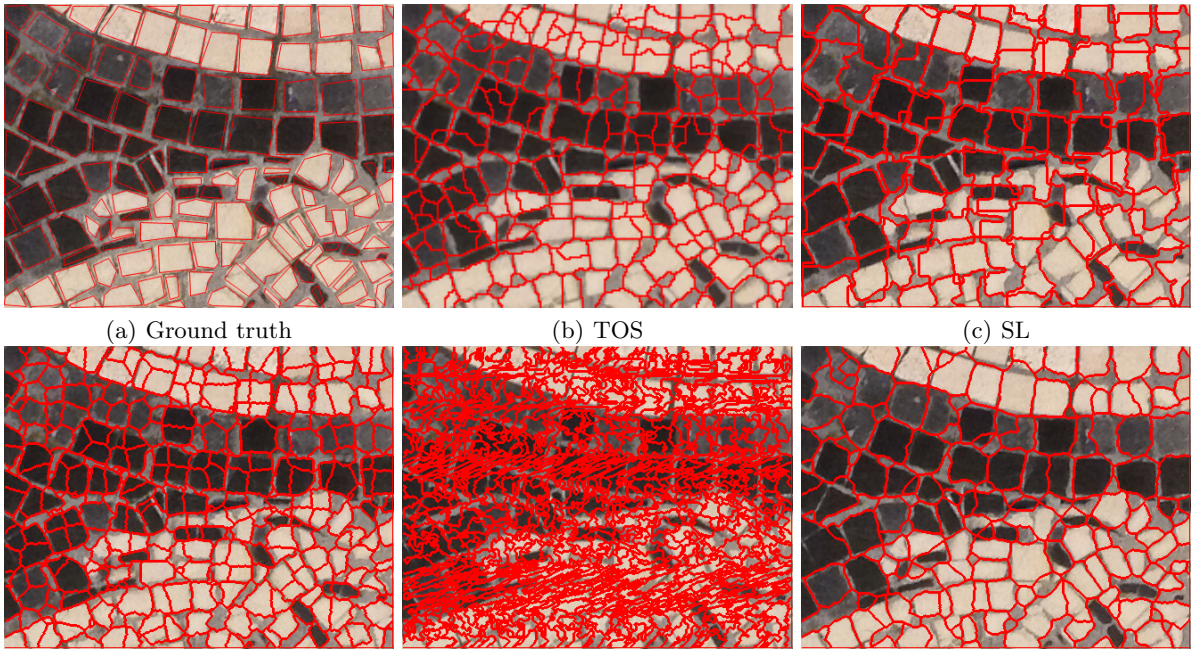
Finally, it can be noted that ERSS and TP methods also perform relatively well on University image, whereas they perform poorly on the Museum image. We think that this can be explained by the fact that the latter image has a higher resolution and these methods are probably not robust w.r.t. the input image resolution.

4 Concluding remarks and future work

We considered the problem of the automatic segmentation of mosaic images. We proposed a set of simple performance metrics and some mosaic-specific hardness estimators. We assessed both on a set of recent segmentation methods which we applied to 5 manually annotated mosaic images. Our experimental analysis shows that:(i) methods which are not specifically tailored to the segmentation of mosaic images are hampered by not taking into account the filler, which introduces a trade-off between F-measure and count error; (ii) in-tile color dissimilarity and out-tile color dissimilarity appear to be good estimators for the hardness of the segmentation of a mosaic image; and (iii) the performance, in terms of F-measure, of the only method which is specifically tailored to mosaic images appear not significantly better than those of other, general-purpose segmentation methods.

Acknowledgements

This work has been funded by FRA 2012, Finanziamento d'Ateneo per progetti di ricerca scientifica, of University of Trieste. The authors would like to thank the Regional Archaeological Service (Soprintendenza per i Beni Archeologici del Friuli Venezia Giulia) for assistance



(a) Ground truth (b) TOS (c) SL

(d) TP (e) ERSS (f) SLIC

Figure 3: Examples of the segmentation obtained for the University mosaic image. The red superimposed lattice represents the boundaries of the found regions.

during the photographic collection of mosaics in Aquileia. The authors would like to thank also Prof. Stéphane Derrode of Ecole Centrale Lyon, LIRIS, for providing the original code of (Benyoussef and Derrode, 2008).

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Susstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(11):2274–2282.

Benyoussef, L. and Derrode, S. (2008). Tessella-oriented segmentation and guidelines estimation of ancient mosaic images. *Journal of Electronic Imaging*, 17(4):043014–043014.

Benyoussef, L. and Derrode, S. (2011). Analysis of ancient mosaic images for dedicated applications. In Stanco, F., Battiato, S., and Gallo, G., editors, *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*. CRC Press.

Levinshtein, A., Stere, A., Kutulakos, K. N., Fleet, D. J., Dickinson, S. J., and Siddiqi, K. (2009). Turbopixels: Fast superpixels using geometric flows. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(12):2290–2297.

Liu, M.-Y., Tuzel, O., Ramalingam, S., and Chellappa, R. (2011). Entropy rate superpixel segmentation. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 2097–2104. IEEE.

Moore, A. P., Prince, S., Warrell, J., Mohammed, U., and Jones, G. (2008). Superpixel lattices. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.

Neubert, P. and Protzel, P. (2012). Superpixel benchmark and comparison. In *Proc. Forum Bildverarbeitung*.

Ren, X. and Malik, J. (2003). Learning a classification model for segmentation. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 10–17. IEEE.

Vincent, L. and Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598.

Yasnoff, W. A., Mui, J. K., and Bacus, J. W. (1977). Error measures for scene segmentation. *Pattern Recognition*, 9(4):217–231.

Zhang, H., Fritts, J. E., and Goldman, S. A. (2008). Image segmentation evaluation: A survey of unsupervised methods. *computer vision and image understanding*, 110(2):260–280.

CHEMICAL REVIEWS

Review

pubs.acs.org/CR

Organic–Inorganic Hybrid Nanocomposite-Based Gas Sensors for Environmental Monitoring

Ajeet Kaushik,*[,†,‡] Rajesh Kumar,*[,‡,§] Sunil K. Arya,[∥] Madhavan Nair,[†] B. D. Malhotra,[⊥] and Shekhar Bhansali[‡]

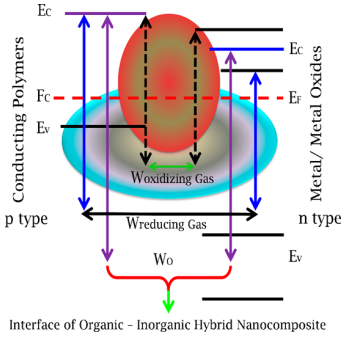
[†]Center for Personalized Nanomedicine, Institute of Neuroimmune Pharmacology, Department of Immunology, Herbert Wertheim College of Medicine, Florida International University, Miami, Florida 33199, United States

[‡]Bio-MEMS Microsystems Laboratory, Department of Electrical and Computer Engineering, College of Engineering, Florida International University, Miami, Florida 33174, United States

[§]Department of Physics, Panjab University, Chandigarh 160014, India

[∥]Bioelectronics Program, Institute of Microelectronics, A*Star, 11 Science Park Road, Singapore Science Park II, Singapore

[⊥]Department of Biotechnology, Delhi Technological University, Shahbad Daulatpur, Delhi 110042, India



Interface of Organic - Inorganic Hybrid Nanocomposite

CONTENTS

1. Introduction A
2. Organic–Inorganic Hybrid Nanocomposites: Smart Advanced Functionalzed Materials for Gas Sensing B
   2.1. Preparation Methods of Organic–Inorganic Nanocomposites C
3. Gas Sensing Principle of Organic–Inorganic Hybrid Nanocomposites E
4. Organic–Inorganic Hybrid Nanocomposite-Based Gas Sensors for Environmental Monitoring F
   4.1. Organic–Inorganic Hybrid Nanocomposites for Volatile Organic Compounds (VOCs) F
   4.2. Organic–Inorganic Hybrid Nanocomposites for Hydrochloric Acid (HCl) Detection K
   4.3. Organic–Inorganic Hybrid Nanocomposites for Ammonia (NH_3) Detection K
   4.4. Organic–Inorganic Hybrid Nanocomposites for Hydrogen Disulfide (H_2S) Detection N
   4.5. Organic–Inorganic Hybrid Nanocomposites for Detection of Nitrogen Oxides (NO_x) N
   4.6. Organic–Inorganic Hybrid Nanocomposites for Carbon Dioxide (CO_2) Detection Q
   4.7. Organic–Inorganic Hybrid Nanocomposites for Carbon Monoxide (CO) Detection R
5. Organic–Inorganic Hybrid Nanocomposites for Humidity Detection S
6. Challenges in Organic–Inorgnic Hybrid Nanocomposite-Based Gas Sensing U
7. Conclusions AB
Author Information AB
   Corresponding Authors AB
   Notes AB
   Biographies AB
Acknowledgments AD
Glossary AD
References AE

1. INTRODUCTION

Rapidly increasing environmental pollution has been recognized as a major concern, and its monitoring has become a priority area for human health.[1−4] This fact has led efforts to find new and user-friendly techniques for the detection of gases hazardous to the environment and human health.[5] Gas sensors (Figure 1) have become one of the key technologies for rapid, selective, sensitive, and efficient detection of gases (at ppt level), chemical vapors (at ∼1 ppb), and explosives (at ∼5 ppm).[6,7] The commercialization of gas sensors for environmental and personalized health monitoring is growing rapidly.[8]

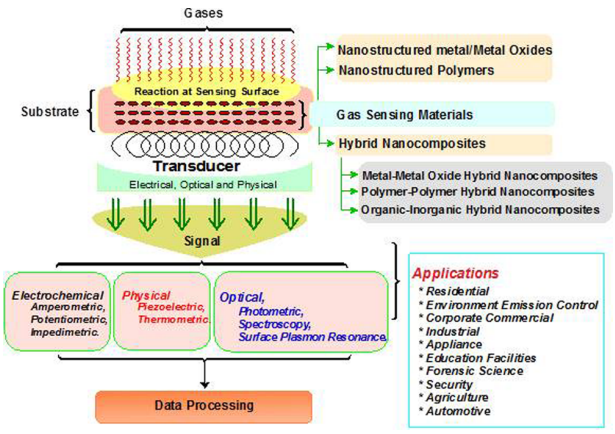


Figure 1. Schematic of gas sensor fabrication techniques.

Received: February 17, 2012

ACS Publications © XXXX American Chemical Society

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

The field of gas sensing device development remains an area of active research interest due to the urgent need for reliable, miniaturized, low-cost, and portable sensing systems for health care and security.[9] The current state of the art in chemical sensors for gas detection has been discussed in reviews by Janata,[10] Korotcenkov,[11] and Stetter.[12]

Gas sensors based on organic conducting polymers (such as polyaniline (PANI),[13,14] polypyrrole (PPy),[15] polythiophene (PTh),[16] etc.) of desired functionality and conductivity continue to improve gas sensing performance.[17−24] However, due to the high affinity of conducting polymers toward volatile organic compounds (VOCs) and moisture present in the environment, they are sometimes unstable and exhibit poor sensitivity.[25,26] On the other hand, gas sensors made from inorganic metal oxides, such as tungsten oxide (WO_3), zinc oxide (ZnO), tin oxide (SnO_2), titanium oxide (TiO_2), iron oxide (Fe_2O_3/Fe_3O_4), silicon oxide (SiO_2), etc., show improved sensing characteristics due to changing oxygen stoichiometry and electrically active surface charge.[27−31] However, these gas sensors usually operate at very high temperatures (~300−400 °C), frequently leading to baseline drift and oxidation of analytes.[32] In spite of the high sensitivity, the use of such gas sensors for practical applications is very limited.

Issues related to low conductivity and poor stability of organic materials, and the high-temperature operation and complicated processability of inorganic materials, hinder their use in gas sensor fabrication. The use of hybrid nanocomposites of these two classes of materials may result in gas sensors with improved and efficient gas sensing characteristics. It has been predicted that host−guest chemistry combined with the use of organic and inorganic counterparts in hybrid nanocomposite forms may help to eliminate their particular drawbacks due to synergetic/complementary effects, leading to development of improved gas sensing devices. These nanocomposite materials have recently aroused extensive interest in gas sensing applications. This review focuses on organic−inorganic hybrid nanocomposite-based gas sensors for environmental monitoring. Additionally, efforts have been made to discuss the gas sensing mechanism of organic−inorganic hybrid nanocomposites and humidity sensing applications. However, better control is required for commercialization of nanostructure-based gas sensors.[33,34]

2. ORGANIC−INORGANIC HYBRID NANOCOMPOSITES: SMART ADVANCED FUNCTIONALZED MATERIALS FOR GAS SENSING

The field of organic−inorganic hybrid nanocomposites is a rapidly growing area of research in advanced functional materials science (Figure 2). These materials are made of nanoscaled organic and inorganic counterparts, wherein the interaction at molecular level generates unique properties at the interface. In nanocomposites, inorganic nanostructures can be three-dimensional, two-dimensional, one-dimensional, or zero-dimensional nanomaterials. Since 1995, the research group of Sanchez has been exploring the origin, fundamentals, advancements, prospects, and possible applications of organic−inorganic hybrid nanocomposites.[35−47] This group has explored transition-metal-based hybrid organic−inorganic structures and proposed this area of research as a multidisciplinary field.[35,36] Design of organic−inorganic hybrid nanocomposites using fundamental nanobuilding blocks (NBB), studies of their properties (i.e., optical, electrical, mechanical, and electrical) along with their nanophotonics application have also been

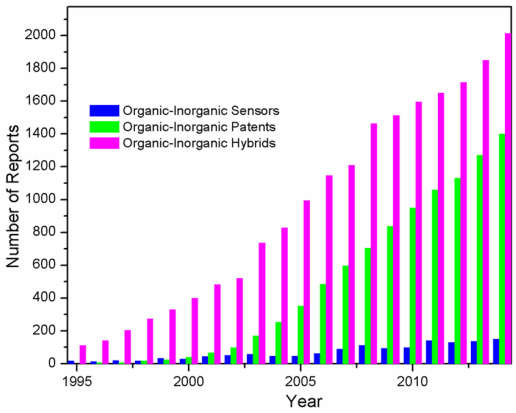


Figure 2. Yearly publications and patents in the field of organic−inorganic hybrid nanocomposites. (Source: web of science and world intellectual property organization.)

explained.[38,39,48−50] The controlled and improved properties generated at the interfaces via molecular and supramolecular dynamics of hybrid organic−inorganic nanocomposites and formation of various mesoporous hybrid nanocomposites have been studied by this research group.[41,42,47,51] In 2005, the various applications of hybrid organic−inorganic nanocomposites were reported by Sanchez et al.[43] These authors proposed a number of organic−inorganic composite systems and advanced hybrid platforms along with their prospects from laboratory to market.[45] The molecular and surface engineering of organic−inorganic hybrid nanocomposite systems in order to improve the properties needed for various applications has also been highlighted by this research group.[44−47]

The properties of hybrid nanocomposite materials depend not only on the properties of their individual constituents, but also on their morphology and interfacial characteristics.[52−56] Comprehensive studies of organometallic compounds have led to increased interest in understanding the interactions of organic and inorganic moieties, and have generated unique properties at interfaces. Further, heterogeneous or homogeneous mixing, suspension, and cross-linking of organic and inorganic components produce interesting composite materials with commercial prospects. Significant efforts have been focused on the ability to obtain nanostructures of desired shape, size, and properties using innovative synthesis approaches.[57] The surface modification and functionalization of nanoparticles, covalent attachment, self-assembly, and ease of organization on the surfaces provide a means to generate nanocomposite materials with tunable surface properties.[52,57] Efforts are continuing toward the preparation of new nanocomposite materials to obtain novel properties.[53,57−59] There is a strong interest in using nanometal or metal oxides as additives in the functionalization of polymers (conducting/biopolymer), and significant research has been undertaken in nanocomposite science for device applications.[60,61] Experimental studies have shown that incorporation of inorganic nanomaterials into polymer matrices leads to organic−inorganic nanocomposites with enhanced and improved properties. Nanocomposites offer diverse properties that are absent in their precursors materials, e.g., flexible-mechanical and superconducting properties due to interfacial interactions between the two phases.[39,62−68] Therefore, nanocomposites have been predicted to have a wide range of applications, e.g., high-energy radiation shielding materials, microwave absorbers,

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

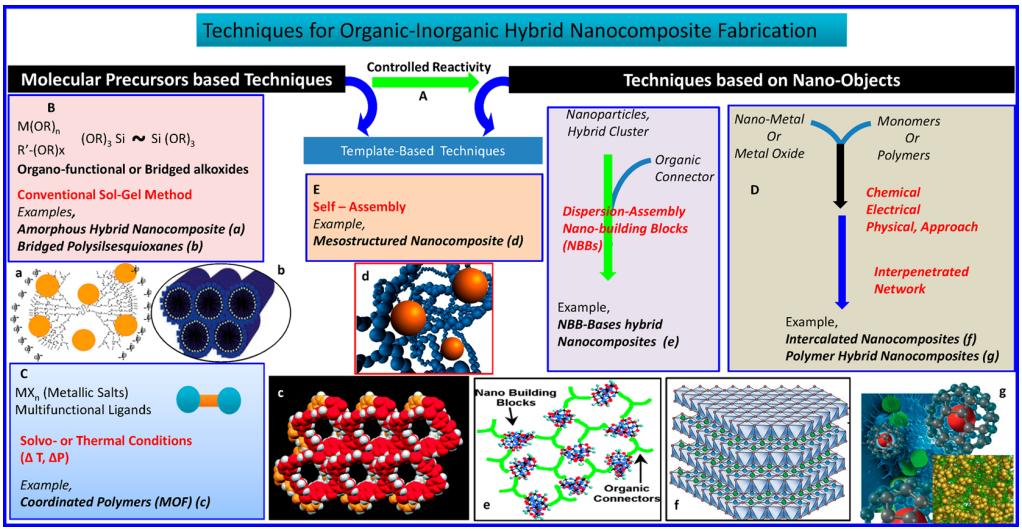


Figure 3. Various strategies adopted for the synthesis of organic–inorganic hybrid nanocomposites along with schematic representation.

optical limiters, polarizers, and hydrogen storage systems.[57,63−68]

A promising and demanding area of research relates to the utilization of organic–inorganic hybrid nanocomposites as gas sensing materials.[69,70] Many reports pertaining to adsorption of gas molecules onto mesoporous ordered organic–inorganic nanocomposite frameworks have been explored since 1990.[68] The effects of parameters such as surface area, pore volume, and pore size on the properties of these materials are still a matter of considerable interest, in order to achieve selective and sensitive detection. Other than interfacial interactions, control of nanostructure morphology is an important aspect that influences the nanocomposite performance. Hence, in a nanocomposite development effort, a long-term goal is to develop new synthetic strategies to control the size, shape, morphology, and composition. This control allows the tuning of nanocomposite properties.[61] The adopted methods to synthesize hybrid nanocomposites of desired properties and application are discussed below.

2.1. Preparation Methods of Organic–Inorganic Nanocomposites

Depending on the nature of the metal or metal oxide and polymer structure, these processes are likely to result in the incorporation of inorganic nanomoieties into the polymer backbone via different electrostatic interactions.[61] Recently, it has been demonstrated that it is possible to develop suitable techniques for these materials, which are often sufficient to yield high-quality nanocomposite films.[61,63,65] The hybrid nanocomposite thin films containing metal or semiconductor nanocrystals and polymers exhibit porous structures.[64] This enables gaseous molecules from the environment to readily penetrate into the polymer matrix and be adsorbed onto the nanocrystals of the nanocomposite film, resulting in strong sensing behavior.[52,58,61,68]

Many techniques such as sol–gel, self-assembly, evaporation deposition, chemical and electrical techniques can be used to fabricate organic–inorganic hybrid nanocomposite thin films.[71−74] Sanchez et al.[40] have described in detail the various strategies (A–E, Figure 3) to synthesize organic–inorganic hybrid nanocomposites of desired architecture and properties for a variety of applications with possible commercial market prospects. Well-defined nonaggregated nano-objects/NBBs were synthesized using path A via controlled reactivity of molecular precursors (metal alkoxides, metal–organic or organometallic compounds, inorganic salts, etc.). The controlled reactivity allows the understanding of nucleation, growth, and aggregation among precursors in water, organic solvents, inorganic molten salts, ionic liquids, gels, xerogels, etc. NBBs are nanoparticles, nanolamellar compounds (clays, layered double hydroxides, lamellar phosphates, oxides, chalcogenides, etc.) with strong molecular interactions. The composition of NBBs can be fully inorganic or hybrid with organic moieties either included in the framework or located at the nano-object surface. For the last two decades, nanosized metals, metal oxides, metal chalcogenides, metal phosphides, and their hybridized associated derivatives with controlled shape, size, and tailored morphology have been synthesized by adopting path A.[39,45,46,51]

Soft chemistry based routes such as sol–gel chemistry have been followed using specific bridged and poly functionalized precursors such as silsesquioxanes $X_3Si\text{-}R'\text{-}SiX_3$ (R′ being an organic spacer, X = Cl, Br, −OR) for the synthesis of hybrid nanocomposites (path B). In this method, "sol" is prepared by dissolving an inorganic precursor in an appropriate solvent (e.g., various alcohols) and controlled addition of water under acidic or basic conditions to initiate the condensation reactions leading to oligomeric molecules or colloids that form a three-dimensional network by chemical reaction.[37,56,73,75−78] The condensation reactions, evaporation of a solvent, and aggregation of a species in the sol result in the formation of a gel that transforms into films on the desired substrate via spin coating, dipping, or spraying techniques. Next, the deposited film is cured at low temperature or annealed. The heat treatment is designed to remove any organic residues and water through evaporation, pyrolysis or combustion, as well as to provide densification. The tunable supramolecular chemical interactions yield hybrid materials with higher degrees of polymerization. For the deposition of an organic–inorganic hybrid film, care must be taken to heat the film below the decomposition point of the organic component.

Path C corresponds to hydrothermal or solvothermal synthesis routes generally performed in polar solvents to

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

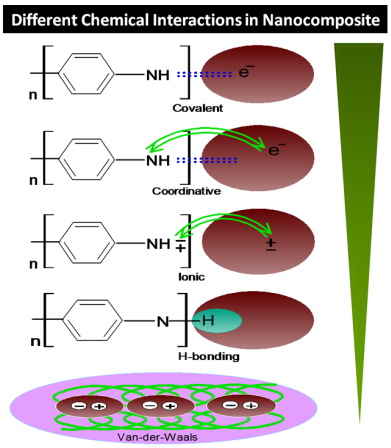
obtain crystallized material. In this method, the utilization of organic templates results in microporous hybrid materials such as organically templated zeolites.[34] These materials have already led to extensive applications in the fields of adsorbents and catalysts.[34] Recently, metal organic frameworks (MOFs), crystalline hybrids of polymers built from telechelic or polyfunctional spacers that coordinate metallic centers or metal-containing small oligomers have been developed.[51]

NBBs predicted to provide nanostructured hybrid networks with better structural definition can be synthesized using assembly or intercalation and/or dispersion (path D). The step-by-step preparation of nanobuilding components that are nanometric, monodispersed, and have well-defined structures and shapes, usually allows for a high level of control over their structure and properties. The NBB molecules retain their integrity in the hybrid nanocomposites, wherein these can be preserved or capped using polymerizable ligands or connected through organic spacers of telechelic molecules, polymers, or functional dendrimers. Different assembling methods were used to synthesize various organic–inorganic hybrids of tunable electrical and optical properties which depend on the nature, structure, and functionality of the nanobuilding blocks.[51]

Path E explores the template-assisted organization and assembly of inorganic or hybrid networks grown using micelles, lyotropic liquid crystals, latex, or silica beads for the fabrication of mesostructured organic–inorganic hybrid composites. In practice, self-assembly of molecular or polymeric amphiphilic components coupled with sol–gel polycondensation reactions or with other colloidal-chemistry-based approaches can be used for growth. Periodically organized mesoporous hybrid silica with organic functionality within the walls was synthesized using path E, and exhibited a high degree of order. The presence of mesoporosity is an added advantage for further organic functionalization. These hybrids can be synthesized by the combination of self-assembly and NBB approaches.[45,46,51] Due to the use of NBBs, a large variety of organic–inorganic hybrid nanocomposites of tunable functionalities and properties can be synthesized via various chemical/molecular interactions (Figure 4). The synthetic approaches discussed above allow control of the shape, size, and properties of the organic–inorganic hybrid nanocomposites. The hybrid nanocomposites are currently being patterned or integrated in devices for applications that demand easy and effective synthesis and processing pathways (Figure 5).

Organic–inorganic hybrid films have recently been prepared via deposition under vacuum conditions, e.g., thermal evaporation and pulse polymerization techniques. The vacuum-deposited hybrid nanocomposite films can be used for fabrication of novel optoelectronic devices due to good repeatability and identical optical, electrical and molecular properties.[52,58,79−82] Organic–inorganic hybrid nanocomposite thin films are fabricated by the aforementioned techniques and can be used as sensing materials for various sensor configurations.[61,83,84] However, the problems pertaining to nanostructure fabrication using expensive sophisticated tools and thermal or temporal incompatibility between the organic and inorganic components during evaporation are still unresolved issues.

Electrical polymerization (Figure 6) is a widely used technique for thin film fabrication of conducting polymer (PANI, PPy, etc.)-metal (Au, CNT, Pt, graphene, etc.)/metal oxide (ZnO, TiO_2, ZrO_2, etc.) hybrid nanocomposites.[85] The heterogeneous electrochemical polymerization of monomers in



Interaction	Strength [kJ/mol]	Range	Characteristics
Covalent	350	Short	Predominantly Irreversible
Coordinative	50-250	Short	Directly
Ionic	50-200	Long	Non-Selective
Hydrogen	5-65	Short	Selective Directional
Van-der-Waals	50	Short	Non-Directional, Non-Directional

Figure 4. Types of chemical/physical interactions during the fabrication of organic–inorganic hybrid nanocomposites.

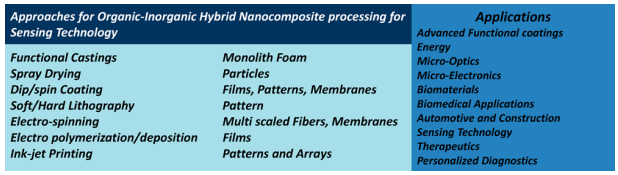


Figure 5. Appropriate approaches for the processing of organic–inorganic hybrid nanocomposites for various applications.

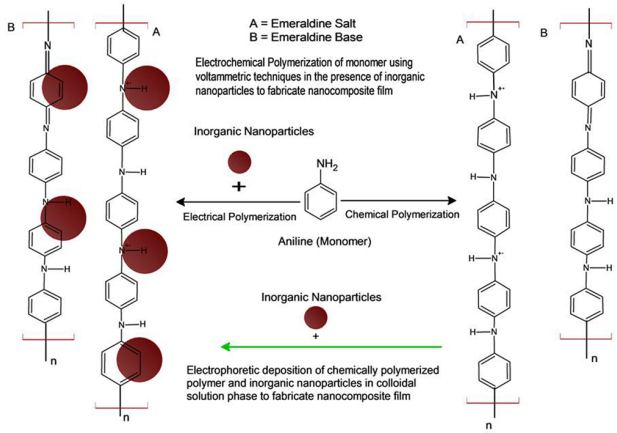


Figure 6. Electrical polymerization and electrophoretic deposition of organic–inorganic hybrid nanocomposites.

the presence of an oxidant, dopant, and metal oxide nanoparticles has been performed using cyclic voltammetry.[86] During polymerization, the presence of metal oxide nanoparticles inhibits the polymer-related oxidation process and diminishes the degradation products in the composite materials. Good adhesion of composite materials on the conducting substrate results in superior electro-active repetitive properties compared to that of a pure polymer. However, the problem pertaining to variation in optical and electrical

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

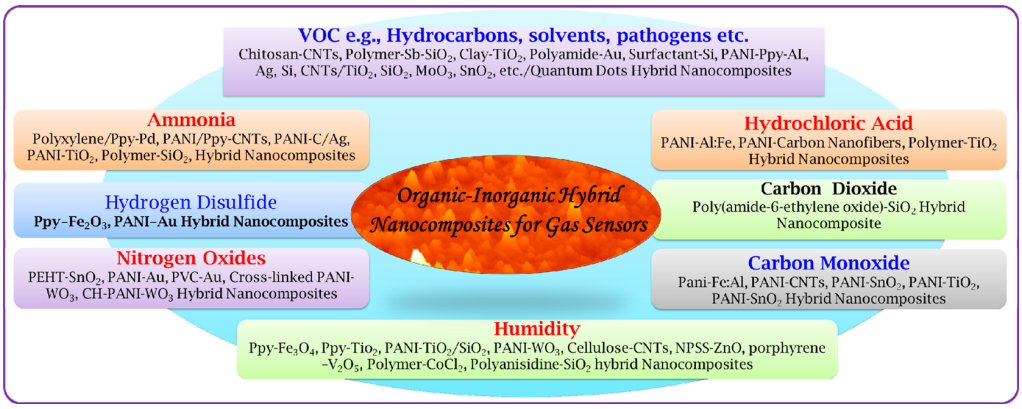


Figure 7. Organic–inorganic hybrid nanocomposites used for gas sensing applications.

properties continues to exist. To overcome this limitation, the electrophoretic deposition technique (two-step process) has been used, wherein the charged particles in a liquid first migrate toward an electrode under an electric field (electrophoresis).[80,87] In the second step, the particles are deposited on the electrode surface from a homogeneous film via coagulation. The electrophoretic technique has recently gained prominence. It has been found to yield uniform, dense, and pinhole-free conducting polymer–metal/metal-oxide-based nanocomposite films. Electrochemically deposited nanocomposite films have been utilized for fabrication of organic light-emitting diodes, display devices, artificial fingerprint devices, photovoltaic applications, sensors, etc.[88,89]

The preparation of a hybrid nanocomposite sensing material with desired properties is dependent on the selection of the synthesis method. Thus, selection of an appropriate preparation technique is crucial to prepare a 3S, i.e., a stable, selective, and sensitive gas sensor. Optical,[90−94] field effect transistor (FET),[95,96] surface acoustic wave (SAW),[97−99] quartz crystal microbalance (QCM),[100] and electrochemical[92,101−104] (i.e., chemoresistor,[105,106] potentiometric,[107,108] and amperometric[12,109]) based gas sensors have contributed significantly to the state-of-the-art of environmental monitoring. The selection of sensor fabrication method and transduction technique is crucial to achieve desired sensing abilities. For example, electrochemical gas sensing techniques are suitable for miniaturized and wearable sensor fabrication.

It has been found that the electrical, optical, and molecular properties of organic–inorganic hybrid nanocomposites can be changed or tuned using an appropriate synthetic route and varying precursors/operational parameters. Moreover, these nanostructures have potentially been used for various technological applications. We herein focus on the gas sensing applications for environmental monitoring. Figure 7 illustrates the organic–inorganic hybrid nanocomposites used to detect gases. The next section focuses on the gas sensing mechanism of the organic–inorganic hybrid nanocomposites.

3. GAS SENSING PRINCIPLE OF ORGANIC–INORGANIC HYBRID NANOCOMPOSITES

The gas sensing behavior of organic–inorganic hybrid nanocomposite sensors obtained by estimating time dependent change in the magnitude of resistance as a function of gas concentration at room temperature is illustrated in Figure 8. It

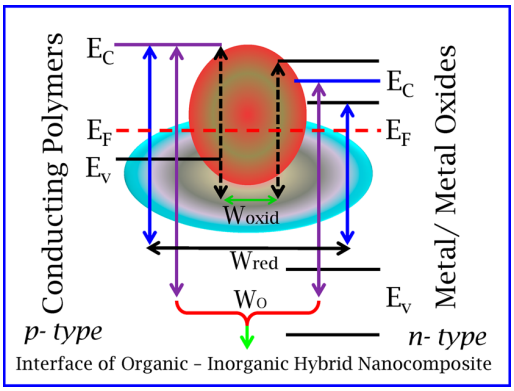


Figure 8. Energy band gap diagram of organic–inorganic hybrid nanocomposites on gas exposure.

has been found that exposure of reducing/oxidizing gas molecules onto nanocomposite film results in a change of resistance. Simultaneously, gas flow is cut off and fresh air is introduced into a test chamber, thus leading to restoration of the original resistance value.[52] The observed change in the resistance is due to physical adsorption of the gaseous molecules onto the surface of the nanocomposite film. The interaction of gaseous molecules with a π electron network of a polymer that has embedded metal/metal oxide nanoparticles results in capture/donation of electrons, depending upon the nature of gaseous molecules, leading to the decrease/increase of the resistance.

The n-type metal oxide nanoparticles form a barrier layer with a polymer matrix leading to the formation of the depletion region. The interaction of the nanocomposite film surface with target gas molecules causes a change in the width of the depletion region (W_0) and therefore modulates the conductivity of the sensing element. The space charge region (W_{gas}) varies with the decrease or increase of electrons (by adsorption of gas molecules) in the polymers, which in turn results in a conductivity change.[52] The modulation of the space charge region at the interface of the metal oxide nanoparticles with the polymer matrix results in enhanced sensitivity for the desired gas. On the basis of the energy band gap diagram (Figure 8), the gas sensing behavior of organic–inorganic nanocomposites can be understood with respect to the analyte.[52,110] The next section focuses on the application of

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

these hybrid nanocomposites to detect toxic gases for environmental monitoring.

4. ORGANIC–INORGANIC HYBRID NANOCOMPOSITE-BASED GAS SENSORS FOR ENVIRONMENTAL MONITORING

The gas sensing potential of organic–inorganic hybrid nanocomposite sensors is discussed in this section. Advancement in sensing materials science and improvement in gas sensing parameters, such as detection limit, sensing range, response/recovery time, repeatability, and stability, are also briefly and critically discussed. Sensor fabrication using organic–inorganic systems for commercial prospects to improve environmental monitoring are also highlighted.

4.1. Organic–Inorganic Hybrid Nanocomposites for Volatile Organic Compounds (VOCs)

The evaluation of indoor air quality is presently a widespread environmental issue.[111,112] Indoor pollutants consisting of VOCs may cause environmental illness.[113,114] The effect of VOCs on human health was brought to the fore during 2007–2009, due to illnesses caused by poorly made residential/commercial dry walls.[115−118] However, the presence of low concentrations and the large diversity of VOCs has made their detection a challenging task.[2,3] Thus, there is increased demand for the development of a continuous real-time technique to monitor indoor VOCs. Recently, nanocomposites fabricated using electrochemical or self-assembly of metal/metal oxide nanoparticles on top of functionalized conducting polymer film surfaces, has enabled the selective sensing of VOCs.[119] In these hybrid nanocomposites, the selectivity is achieved by assembling different metal nanoparticles on the same conducting polymer film. This eliminates the need to develop different polymer chemistries and device configurations for each specific analyte. The utilization of organic–inorganic nanocomposite-based VOC sensors are described briefly in this section. The aspects of nanocomposite material science resulting in sensing performance and the related sensing mechanism are also discussed.

For the detection of hydrocarbon vapors, Li et al.[120] used a PANI-MWCNT (mass ratio 4:1) hybrid nanocomposite prepared by an in situ electrochemical polymerization technique. This sensor exhibited a response in the 200–1000 ppm range of aromatic hydrocarbon vapors due to an increase in conductivity, and the maximum response (0.31%) was measured at 1000 ppm. The increase in conductivity of PANI after gas exposure has been attributed to physical interactions due to dipole–dipole interactions that uncoil the polymer chain and decrease the hopping distance for the charge carriers. Electrochemically fabricated PANI/aluminum (Al) nanocomposite-based Schottky barrier diodes have been used by Campos et al.[121] to detect methanol (CH_3OH). The current–voltage (I–V) and capacitance–V characteristics of the diodes were studied with different concentrations of gas for different time intervals at room temperature. The presence of CH_3OH resulted in a decrease in the forward and reverse-bias current and an increase in the barrier height, i.e., increase in resistance. However, the effects of different concentrations and sensitivity were not explained and thus need more investigation. An optical CH_3OH sensor has been fabricated by Stevens et al.[93] using a mesoporous coumarin 481 (C481) dye-doped hybrid composite film of nonionic surfactant Pluronic P123 and silica by sol–gel methods. The composite films were found to be capable of CH_3OH detection at a concentration of ∼150 ppm via changes in the fluorescence intensity. An energy transfer complex formed via interaction between dye and nanocomposite exhibited enhanced response to methanol vapor, but only for the first few sensing cycles. The observed decrease in fluorescence intensity is attributed to an excited state interaction between methanol and C441. However, key parameters of these films, such as sensitivity, long-term photostability, and specificity to methanol, need further optimization to fabricate commercially viable optical sensors. Sadek et al.[122] described a surface acoustic wave (SAW)-based gas sensor to detect H_2 using a nanocomposite film of PANI-WO_3 nanofiber fabricated onto a layered ZnO/64°YXLiNbO_3 SAW transducer via chemical oxidative polymerization. It was observed that, during sensing, the analyte adsorbed onto the WO_3 surface and generated free electrons, resulting in a decreased depletion region and leading to increased surface conductivity. The sensor response was found to be 7 kHz toward 1% of H_2 in synthetic air at room temperature with a response time of 40 s, recovery time of 100 s, and repeatability in a stable baseline condition. A hybrid nanocomposite of plasma-polymerized acrylonitrile (PP-AN)-ZnO nanobelt (NB) has been used to detect O_2, and its sensing characteristics have been compared with bare ZnO nanobelt based sensors.[123] The nanocomposite showed better O_2 sensing than bare ZnO-NB. The obtained sensing performance enhances on UV exposure. UV exposure increases the conduction in the nanocomposite due to generation of a photocurrent, resulting in a direct increase of electron–hole pairs. Moreover, desorption of charge species with a related reduction in the electron depletion region near the ZnO-NB surface leads to modified surface potential. It was observed that nanocomposite-based sensors exhibited response enhancements from 0 to 25.3% (minimum sensitivity at 150 °C is 16.6 ppm in 300 s) at low temperature. Also, the sensor can avoid structural deterioration and caused stability in the obtained response. However, the effects of O_2 desorption/adsorption in PP-AN on the electron depletion region in the ZnO were significant for sensing phenomena.

Solvents such as propanol ($CH_3CH_2CH_3OH$), toluene ($C_6H_5CH_3$), ethanol (CH_3CH_2OH), hexane, (C_6H_{14}) and chloroform ($CHCl_3$) were detected using array-based gas sensors using hybrid composite films of polyvinyl butyrl (PVB), Fe_2O_3 and manganese dioxide (MnO_2) by Arshak et al.[124] The response of the sensors $\Delta R = \{(R_{gas} - R_{air})/R_{gas}\} \times 100$ (R_{gas} and R_{air} are the magnitude of resistances in the presence of target gas and air, respectively) was seen to increase with increasing gas vapor concentration and was found to be dependent on film compositions and the number of layers. The PVB containing 1.5 wt % carbon black exhibited a response to CH_3CH_2OH (43.27%) and $C_6H_5CH_3$ (19.43%) at 5000 ppm. In contrast, the response of a sensing layer comprised of Fe_2O_3/MnO_2 to the same concentration of $CH_3CH_2CH_3OH$ and $C_6H_5CH_3$ at room temperature was found to be 62.57% and 45.95%, respectively. Further, authors investigated the application of nickel oxide (NiO)/Fe_2O_3 (75/25, 50/50, and 25/75) and PVB nanocomposite films for gas sensor applications to detect CH_3OH, CH_3CH_2OH, CH3CH_2CH_2OH, $C_6H_5CH_3$, acetone (CH_3COCH_3), and tetrahydrofuran (THF, C_4H_8O) at room temperature (25 °C).[125] For each composition tested, the highest response was displayed for $C_6H_5CH_3$ and CH_3CH_2OH. The fastest response and recovery times of each device were recorded after their exposure to 4000 ppm of $C_6H_5CH_3$ and CH_3CH_2OH at room

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

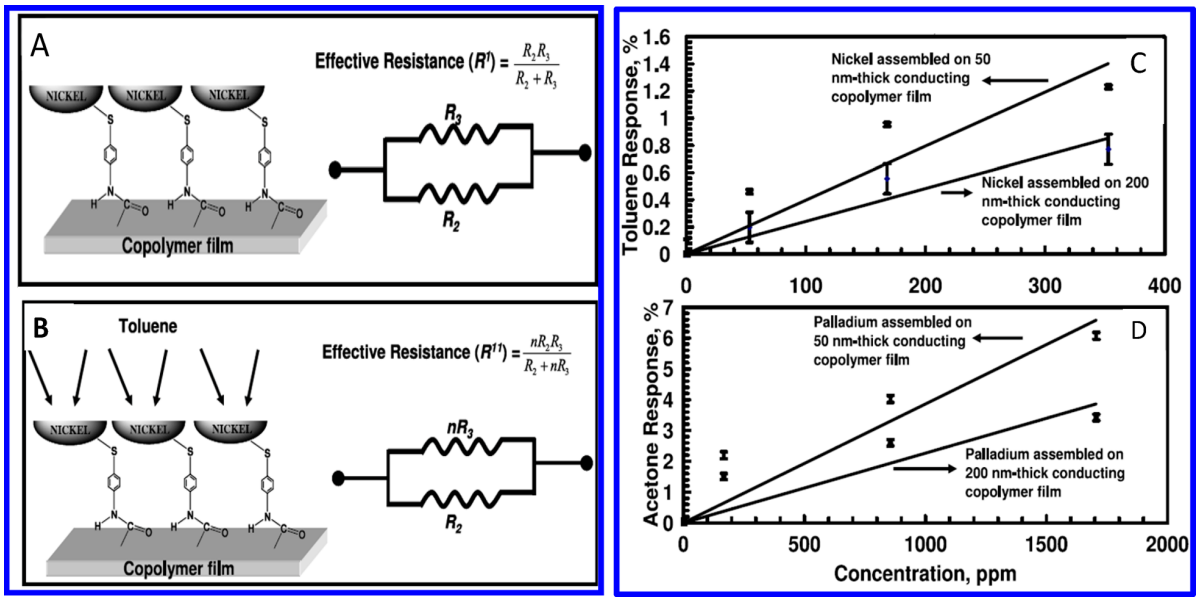


Figure 9. Pictorial representation of Ni-poly(EDT-*co*-TAA) before and after exposure to toluene (A and B); variation of the response of Ni-poly(EDT-*co*-TAA) hybrid toluene concentration (C); and (D) variation of the response of Pd-poly(EDT-*co*-TAA) hybrids to acetone concentration. Reprinted with permission from ref 126. Copyright 2010 Institute of Physics.

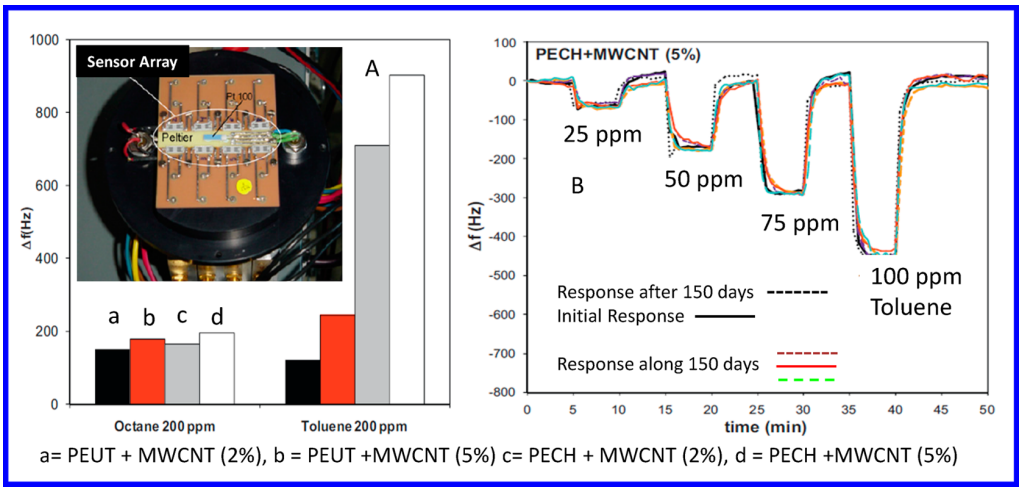


Figure 10. Inset in (A), the sensor array in the measuring cell and the oscillator circuits, (A) sensing responses at 200 ppm of octane and toluene at room temperature. (B) Testing of repeatability of the S4 sensors exposed to different concentrations of toluene over 5 months at room temperature. Reprinted with permission from ref 127. Copyright 2012 Elsevier.

temperature. Sensors fabricated using composition 75/25 M wt % of NiO to Fe_2O_3 showed response/recovery times of 30/45 s for CH_3CH_2OH and 40/48 s for $C_6H_5CH_3$.

The detection of acetone (CH_3COCH_3) and $C_6H_5CH_3$ at ppm level is achieved by covalent assembly of Ni and Pd metal nanoparticles on top of poly(3,4-ethylenedioxythiophene-*co*-thiophene-3-acetic acid), poly(EDOT-*co*-TAA), films using 4-aminothiophenol linker molecules (Figure 9).[126] These authors explained the effect of Ni/Pd-poly(EDOT-*co*-TAA) hybrid film thickness on the sensitivity to $C_6H_5CH_3$ and CH_3COCH_3. The initial changes in resistance of Pd/poly(EDOT-*co*-TAA) and Ni/[poly(EDOT-*co*-TAA)] hybrid films (50 nm) were at 10 ppm with a response time of 60 s to CH_3COCH_3 and at 2 ppm with a response time of 200 s to $C_6H_5CH_3$, respectively. The observed enhanced performance was attributed to the improved interactions between precursors due to covalent binding. However, Pd/poly(EDOT-*co*-TAA) films were found to be selective for CH_3COCH_3, and Ni/poly(EDOT-*co*-TAA) films were shown to be selective for toluene. The observed sensing was thickness-dependent, and the optimized thickness of the sensor was estimated at ~50 nm.

A SAW gas sensor (Figure 10) based on nanocomposites of MWCNT with polyepichlorohydrin (PECH) and polyetherurethane (PEUT) as sensitive layers was investigated to detect octane (C_8H_{18}) and $C_6H_5CH_3$. The effect of MWCNT concentration on the response of sensors has been studied.[127] It was observed that the introduction of small percentages (~1%) of MWCNT into the polymer increased the adsorption of toluene, but no change has been observed for C_8H_{18} (Figure 10A). The fabricated sensor shows no response on exposure of

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

nitrogen dioxide (NO_2), ammonia (NH_3), H_2, or carbon monoxide (CO), confirming its selectivity. The detection limits of C_8H_{18} and toluene were found to be 7.4 and 1.7 ppm, respectively (Figure 10B). The stability and repeatability of the sensor responses have been confirmed by testing the sensors periodically for 5 months.

Chloroform ($CHCl_3$), one of the most widely used solvents, has been detected using quantum dot (QD) nanocomposite-based gas sensors. QDs of cadmium selenide/zinc sulfide (CdSe/ZnS) were added and encapsulated within polystyrene (PS)−polystyrene-*co*-maleic anhydride (PSMA); PS−PSMA polymer nanofibers using electrospinning techniques to fabricate nanocomposite sensing films to detect chloroform and THF.[128] Change in electrical conductivity due to interactions between the QDs and the PSMA side-chains were used to detect organic solvents with a response time of less than 1 s and with high durability. The fabricated sensor exhibited a shelf life of 6 months and was found to be more sensitive to chloroform (0.001% to 0.99%) than to THF. Olive oil capped QDs of CdSe (~6 nm) and poly(3-hexylthiophene) nanocomposite sensors have been used for selective detection of chloroform at room temperature.[129] This nanocomposite sensor exhibited a linear range from 100 to 1200 ppm and a detection limit of 100 ppm. The sensor was found to be selective to chloroform vapors in comparison to VOCs such as toluene, propanol, and acetone with concentrations of 1000 ppm. The recovery time of the sensor improved from 8 to 4 min on illumination with monochromatic light (600 nm), due to photoinduced enhancement of charge transfer in the nanocomposite under the incident light.[129] Sharma et al.[130] reported Cu-PANI nanocomposite-based gas sensors for chloroform detection at the ppm level, wherein the response time in terms of increase in DC electric resistance on exposure to chloroform vapor was observed. In comparison to the pure PANI, the nanocomposite exhibited good, reversible response for a comparatively low concentration of chloroform ($CHCl_3$, ~100 ppm) with a relative sensitivity range of 1.5−3.5. This gas sensing mechanism mainly involves the adsorption−desorption of $CHCl_3$ on a Cu nanocluster surface. Chemisorption of the analyte vapor induces charge redistribution in the metal nanoparticles and changes their work function. The conjugated linker molecule responds to the change in the work function of the embedded nanoparticles and reflects the changes in the electronic states of the underlying conducting polymer films, resulting in a measurable change in the resistance of the hybrid nanocomposite.

Ma et al.[131] deposited nanocomposite thin films of PANI-TiO_2 onto interdigitated carbon paste electrodes via a spin-coating and immersion method for detection of trimethylamine at room temperature. This nanocomposite film exhibited gas-sensitivity to trimethylamine [$N(C_2H_5)_3$, 5.14×10^{-7} mol mL^{-1}], which increases further by 3−5 orders of magnitude within 3 min and was selective to analogous gases. The sensing film exhibited reproducibility, stability, and easy recovery with high-purity N_2 at room temperature. The nanocomposites of polysiloxanes containing 3-heptafluorobutanoyl camphorates of metal ions, e.g., Zn(II), Ni(II), and Eu(Ill), have been used by Hierlemann et al.[132] as sensitive layers for thickness-shear mode resonators (TSMRs) for the monitoring of various VOCs. Sensing layers provided pronounced selectivity in detecting VOCs in the gas phase and in particular analytes containing heteroatoms, such as O and N. The selectivity resulted from coordinative binding between the electronically unsaturated metal complexes and the heteroatom-containing analytes. The detection of analyte in the concentration range 60−100 ng L^{-1} or 20−40 ppb and a detection limit of 1 ng L^{-1} were achieved for pyridine (C_5H_5N) and *n*-butylamine ($CH_3CH_2CH_2CH_2NH_2$). The drawback of long response and recovery times can perhaps be compensated by applying an evaluation method based on the response slope.

Wang et al.[133] reported fabrication of nanocrystalline antimony-doped tin oxide (ATO ~ 5 nm) with an amphiphilic block copolymer [poly(ethylene-*co*-butylene)-*block*-poly-(ethylene oxide] template via the sol−gel method. It was found that polymers not only provide favorable sites for the growth of the particulate assemblies, but they also influence the formation process, including the nucleation, growth, coagulation and flocculation. Formaldehyde (HCHO, 10−50 ppm) gas sensing devices were fabricated using ATO nanoparticles on alumina tubes with Au electrodes and Pt wires, and showed quick response to HCHO gas (0 to 10%) at an operating temperature of 136 °C. This sensor exhibited a sensitivity coefficient of 3.44 and a low detection limit of 130 ppb. However, there is considerable scope to improve the sensing parameters of this sensor. To detect aldehydes at room temperature, Wang et al.[134] explored ion-exchange reactions between PANI and MoO_3 to prepare $(CH_3CH_2CH_3)$-$(PANI)_xMoO_3$ nanocomposite films. The fabricated nanocomposite film exhibited an increase of 8% in electrical resistivity for 500 ppm formaldehyde and a change of 3.8% for 500 ppm acetaldehyde within 600 s at 30 °C. However, no response was observed for toluene and xylene. Authors suggested that, in the nanocomposite, PANI amplified the signal and the analyte affected the charge conduction between the host and guest.

Itoh et al.[135,136] explored layered organic−inorganic hybrid nanocomposite films of molybdenum oxide (MoO_3) with poly(*N*-methylaniline) (PNMA), PANI, and poly(*o*-anisidine) (PoANIS) formed by a modified intercalation process to investigate the effect of aldehyde such as HCHO and acetaldehyde (CH_3CHO). The $(PNMA)_xMoO_3$ hybrid nanocomposite thin film exhibited increased resistivity (S = 2.8%) to approximately 1−10 ppm for aldehyde gases with reproducibility. However, $(PANI)_xMoO_3$ and $(PoANIS)_xMoO_3$ hybrid thin films exhibited enhanced response magnitude (S = 6%) as a function of resistance when exposed to HCHO and CH_3CHO in the range of 25−400 ppb at 30 °C.[136] The obtained high sensitivity and low quantity detection using $(PANI)_xMoO_3$ and $(PoANIS)_xMoO_3$ nanocomposite-based sensors is due to the modified sensor fabrication approach used, which removes insoluble polymer content and enhances interlinking between polymer and MoO_3 nanoparticles. Additionally, the sensitivity of the (PNMA)xMoO_3 hybrid thin film was found to be suitable and selective for CH_3CHO. It was reported that $(PNMA)_xMoO_3$ hybrid composites possess interlayer spaces equal to the size of methyl group, which assist in adsorption of CH_3CHO.[135]

Lu et al.[137] fabricated a layer-by-layer nanocomposite film of PANI nanoparticles and multiwall carbon nanotubes (MWCNT) onto interdigitated electrodes for fabrication of stable chemiresistive sensors for CH_3OH, CH_3CH_2OH, and $C_6H_5CH_3$ detection with reproducible response upon chemical cycling. Double percolated conductive networks in PANI (1%)-MWCNT (0.005%) hybrids resulted in both higher sensitivity (relative amplitude ~1.1%) and selectivity than other formulations, demonstrating a positive synergy. These features

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

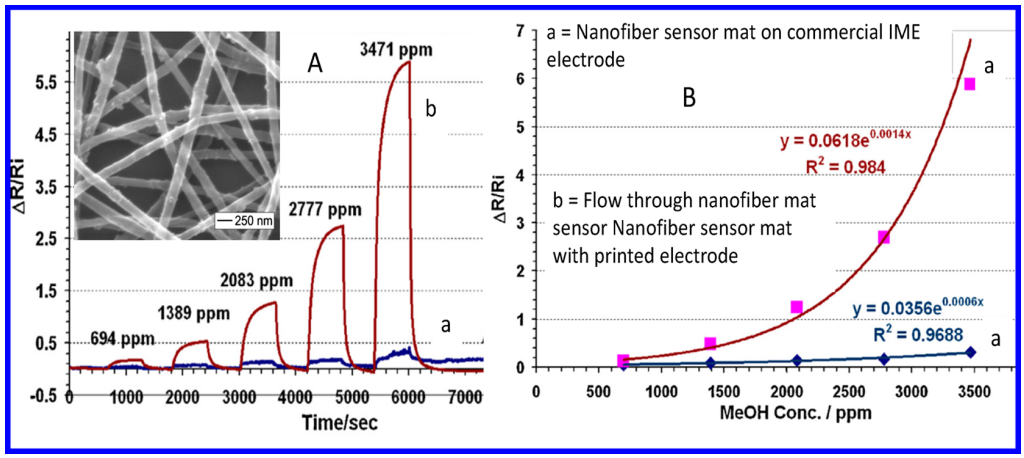


Figure 11. Inset in (A) SEM image of typical electrospun SWCNT-PMMA composite nanofibers. (A) Comparison of the sensor response profile of a flow-through sensing material with a printed electrode and a nonflow-through sensing mat on commercial IME electrodes, and (B) comparison of the sensitivities. Reprinted with permission from ref 143. Copyright 2013 Elsevier.

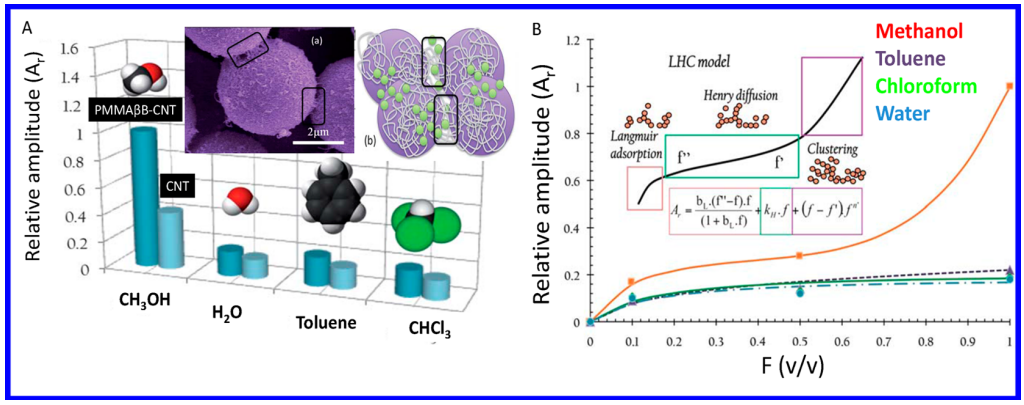


Figure 12. Inset (A) SEM image of MCNT-PMMA hybrid (a), schematic presentation of bridging of PMMA and (b) in a MWCNT random network. (A) Selectivity of a MWCNT-PMMA and CNT random network when exposed to methanol, water, toluene, and chloroform vapors. (B) Effect of analyte concentration on MWCNT-PMMA response amplitude Ar for methanol, water, toluene, and chloroform vapors. Reprinted with permission from ref 144. Copyright 2011 Royal Society of Chemistry.

make them attractive for integration with electronic noses to balance the stability and sensitivity. The present report was mainly focused on the effect of precursor concentrations on sensing characteristics. Hence, efforts should be made to improve selectivity and sensitivity of this sensor.

Optical gas sensors based on Au-polyimide nanocomposite films fabricated by glow discharge vapor deposition polymerization for CH_3OH and CH_3CH_2OH detection using the SPR technique have been reported by Fernandez et al.[138] It was observed that a pure polyimide film does not show any absorption change in the presence of alcohol vapors, but good response was observed in the case of Au-polyimide (5×10^{16} Au^+/cm^2), which exhibited a dynamic optical absorption sensing response for CH_3OH (6000 ppm) and CH_3CH_2OH ethanol (6000 ppm) in the spectral range corresponding to the SPR peak of the Au nanoparticles. The SPR technique has been found more suitable for gas sensing; however, limitations regarding the suitable film thickness for SPR excitation are yet to be overcome. The reported sensitivity is low and shows low selectivity. Nanocomposites of MoO_3 and PPy have been prepared using the vapor polymerization technique for application to CH_3CH_2OH sensing.[139] It was observed that at room temperature the MoO_3−PPy composite exhibited an improved sensing response when compared to that of a MoO_3 or PPy based system. The magnitude of the resistance of the composite was found to decrease with increasing concentration of CH_3CH_2OH (0.3%), and this was due to catalytic reaction of CH_3CH_2OH on the MoO_3 surface. A hybrid nanocomposite of PANI-silver (Ag, ∼15 nm) prepared using ultrasound-assisted in situ mini-emulsion polymerization has been used for ethanol sensing by recording changes in electrical conductivity.[140] The sensing ability and response time of the PANI-Ag nanocomposite-based sensor was found to be dependent on addition of the Ag nanoparticles into the polymer matrix. This nanocomposite exhibited a detection range of 50−225 ppm, a response time of 2100 s, and a maximum response at 100 ppm.

For CH_3OH sensing, the layer-by-layer technique has been used by Kumar et al.[141] to fabricate conductive bionanocomposites of chitosan (CH)-CNT with good chemical and electrical properties. CH−CNT nanocomposites exhibited a chemical nanoswitching mechanism in promoting tunneling conduction. Using this bionanocomposite, the authors described the chemical and electronic behavior of the nanocomposite after diffusion of a gas molecule, resulting in

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

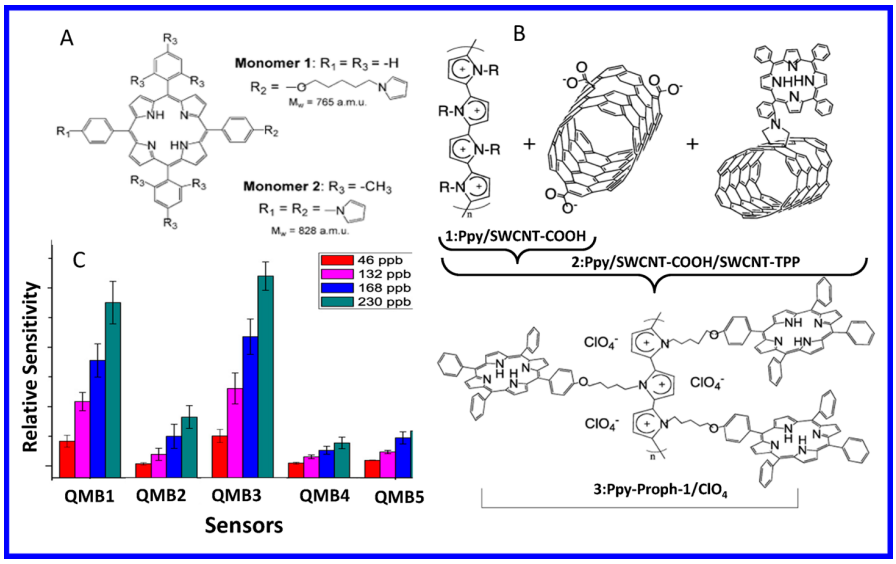


Figure 13. (A and B) Chemical structure of TPP and Pr monomers 1 and 2, schematic representation of QMB coatings, where R in PPy backbone corresponds to H or monomer 2 substituents, and the relative sensitivity to 1-butanol for QMBs modified with SWCNT−Pr electro-polymers. (C) Relative response study of all 5 QCM gas sensors as a function of gas concentration. Reprinted with permission from ref 145. Copyright 2012 Elsevier.

enhanced tunneling conductance. This hybrid bionanocomposite-based transducer showed high sensitivity toward water and CH_3OH (0 to 100%) and showed low response to $C_6H_5CH_3$. Athawale et al.[142] prepared a PANI-palladium (Pd) nanocomposite using oxidative polymerization to fabricate a chemiresistive gas sensor to detect different aliphatic alcohol vapors such as CH_3OH, CH_3CH_2OH, and isopropanol [$(CH_3)_2CHOH$]. Due to its high polarity, the fabricated sensor showed higher selectivity and sensitivity (8.9×10^5 Ω ppm^{-1}) to CH_3OH (10 ppm) vapors than that of CH_3CH_2OH and $(CH_3)_2CHOH$, which causes more efficient interaction on sensing surfaces. However, these sensors exhibited high response time (~2 s) due to the competition arising between the different molecules adsorbed over the nanocomposite surface, wherein the presence of bulkier molecules resulted in a barrier diffusing toward the sensor, compared with the smaller CH_3OH molecules. A flow-through sensor was designed by Han et al.[143] Authors fabricated a SWCNT-poly(methyl methacrylate) (PMMA) nanofiber hybrid composite onto IDE using electrospinning techniques, which was used as a flexible sensing material to detect CH_3OH (600−3500 ppm). Ag contacts printed on nanofiber structures using printing techniques were used for measurements. It was observed that this flexible flow-through sensor design shows 1 order of magnitude higher sensitivity than that of a nanofiber deposited on a conventional IDE (Figure 11).

A 3D sensing platform has been developed by Feller et al.[144] via assembling MWCNT onto PMMA microbeads using spray-coated layer-by-layer (LBL) techniques to fabricate a chemiresistive VOC (CH_3OH, H_2O, $C_6H_5CH_3$, and $CHCl_3$) sensor.[144] It was found that, in this hybrid nanocomposite, the system had an original hierarchical architecture with a segregated network of MWCNT bridging PMMA (Figure 12, inset A). The sensors showed an increase in resistance when exposed to vapor stream, i.e., with target gases adsorption, and also showed a decrease in resistance when exposed to dry nitrogen, i.e., desorption of analyte. This behavior was attributed to the fact that the diffusion of solvent molecules results in disconnection of the conducting network of CNTs, because the energy required for electron circulation by tunneling increases exponentially with the CNT-CNT junction gap. The MWCNT-PMMA hybrid nanocomposite exhibited selective response on exposure to CH_3OH (Figure 12A). The effect of analyte concentration on MWCNT-PMMA response amplitude for CH_3OH, H_2O, $C_6H_5CH_3$, and $CHCl_3$ vapors were studied (Figure 12A), and the curves were fitted with the Langmuir−Henry clustering (LHC) electro-sorption model described in the inset of Figure 12 B. The outcome of sensing results showed that the sensitivity of 3D assemblies of the MWCNT network increased by a factor of 2 and selectivity for CH_3OH vapor by a factor of 5 relative to other VOCs. This sensor can be used as an "electronic nose", which requires a balance between stability and sensitivity.

To detect 1-butanol, Lvova et al.[145] utilized SWCNTs modified with porphyrin (Pr) hybrid composites as sensing materials in quartz microbalance (QMB) sensors with an integrated preconcentration unit. A CNT-Pr hybrid composite was fabricated following two different strategies (Figure 13). Two methodologies were adopted to prepare QCM sensing materials. First, the electropolymerization of pyrrole-substituted Pr monomers, wherein negatively charged carboxylic acid (COOH)-functionalized SWCNTs acted as a dopant. Second, SWCNT−Pr composites were synthesized chemically and then entrapped in PPy/SWCNT-COOH films. SWCNT-Pr was utilized both as the preconcentrator adsorbing phase and QMB coating for sensitive detection of 1-butanol ($CH_3CH_2CH_2CH_2OH$), in concentrations close to the odor threshold. Five different QMB sensors were fabricated using preconcentration procedures to detect low concentrations of $CH_3CH_2CH_2CH_2OH$. QMB 1 (PPy/SWCNT-COOH), QMB 2 (PPy/PPy-Porph-2/SWCNT-COOH), and QMB 3 (PPy/SWCNT-COOH/SWCNT-TPP) exhibited higher responses than QMB 4 (PPy-Porph-1/ClO_4) and QMB 5 (PPy-Porph-2/ClO_4) due to the presence of SWCNTs (Figure 13). The

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

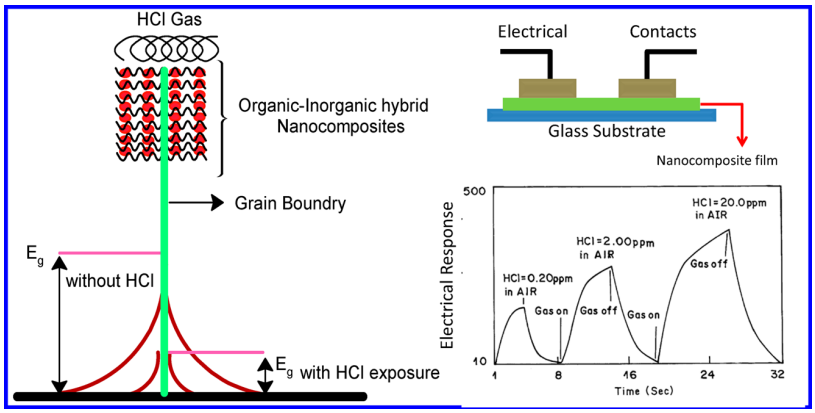


Figure 14. Effect of HCl gas on a organic−inorganic hybrid nanocomposite, schematic of gas sensor fabrication, and gas sensing response of a nanocomposite as a function of HCl concentration. Reprinted with permission from ref 82. Copyright 2004 Elsevier.

results exhibited detection of 1-butanol from 46 to 500 ppb using the QMB 3 sensor.

4.2. Organic−Inorganic Hybrid Nanocomposites for Hydrochloric Acid (HCl) Detection

Hydrochloric acid (HCl) is known to be a level III hazardous air pollutant. HCl is the most commonly used acid in the semiconductor and chemical industries.[146,147] It is excreted in incineration plants, chlorofluorocarbon (CFC) atmospheres, and accidental emissions. HCl is a source of the well-documented pollutants acid rain and dioxin, and has also been identified as a workplace hazard with a short-term exposure limit of 5 ppm. The toxic nature of HCl in both gaseous and aqueous forms has led to increased demand for detection of HCl at ppm levels.[1] It is therefore essential to fabricate sensors to monitor the presence of HCl at site using a fast, user-friendly, safe, and economical method.[146,147] Recently, there has been remarkable growth in the development of electro-active semiconducting polymer thin-film-based sensors for detection of HCl gas.

Nanocomposite polymeric combinations have been found to play an important role in the development of efficient sensors for detection of HCl and other hazardous gases. Jang et al.[148] fabricated carbon nanofiber (CNF ~50 nm)-PPy (~22 nm) coaxial nanocables (~75 nm) via one-step vapor deposition polymerization (VDP) to detect ammonia (NH_3) and HCl. The PPy-coated CNF exhibited an enhanced response signal (from 1 to 3.5, estimated using $\Delta R/R_o$) due to the presence of the thin and uniform conducting polymer layer. The observed response was found to be dependent on the PPy layer thickness. The electrical conductivity of the sensor was found to decrease upon NH_3 exposure due to an increase in the fraction of neutral polymer chains through electron donation to oxidize the polymer chain. The conductivity increased after HCl exposure because Cl^- ions embedded in the polymer backbone lead to the formation of polarons/bipolarons. The response of the PPy-coated CNF upon interaction with NH_3 (10 ppm) and HCl (20 ppm) was reversible and reproducible. However, there is scope to improve the selectivity, sensitivity, and detection limit of these materials. An optically transparent nanocomposite thin film of tetracationic porphyrin (TMPyP) and TiO_2, fabricated using glancing-angle physical vapor deposition techniques, exhibited direct monitoring of HCl.[149] The fabricated sensor had a recovery period of 300 s, a response time of 50.7 s, a detection limit (0.1 ppm), and a shelf life of 30 days. The obtained detection of 0.1 ppm is remarkable; however, this fabricated sensor operates at 80 °C.

A selective, rapid and sensitive HCl gas sensor was fabricated by Mishra et al.[82] using nanocomposites of copolymers of aniline and formaldehyde synthesized with a metal complex of Fe−Al (95:05) via thermal vacuum evaporation deposition techniques. This sensor detects HCl (0.2 to 20 ppm) within 8−10 s. These nanocrystalline composite films exhibited high sensitivity (400−800) and a response time of 10 s (Figure 14). During sensing, exposure of HCl reduces the barrier height at the metal−polymer interface, which increases the charge transport across the interface and enhances output current. The selectivity was achieved by suitable doping of PANI during synthesis (Figure 14). The sensor was reusable, as there was no chemical reaction between PANI film and the gas. Furthermore, the sensor operated at room temperature and possessed an extended lifetime.

4.3. Organic−Inorganic Hybrid Nanocomposites for Ammonia (NH_3) Detection

Ammonia (NH_3) is an important gas to be monitored when studying air pollution. High dosages of NH_3 are easy for humans to detect due to the strong odor of the gas. However, small quantities need to be determined using highly sensitive and selective sensors. Some traces of NH_3 found in drinkable water indicate decomposition of organic matter that may be harmful to human health.[150] In contrast, farmers frequently use ammonia-based fertilizers, which at high doses alter the acidity of the soil, creating an ecological imbalance.[151,152] These factors have motivated researchers to expand their efforts for the development of a selective and sensitive gas sensor to monitor NH_3 at room temperature.

Ammonia, along with VOCs, has been detected using an organic−inorganic hybrid nanocomposite based gas sensor. A hybrid nanocomposite of PANI and 3-mercapto-1-propane-sulfonate-functionalized Au nanoparticles has been prepared using a new and facile method based on osmosis. The PANI and Au NPs were dissolved together in dimethylformamide (DMF) in a dialysis bag and immersed in water.[153] The nanocomposite has been tested for sensing of NH_3 and other VOCs. The sensitivity toward H_2O was 0.00055% ppm^{-1}, whereas it was 1.9% ppm^{-1} for NH_3. The doped PANI-Au NPs nanocomposite shows a low detection limit of 10.8 ppm at room temperature. The NH_3 gas sensing properties of five PANI based materials were studied by Lobotka et al.: thick and

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

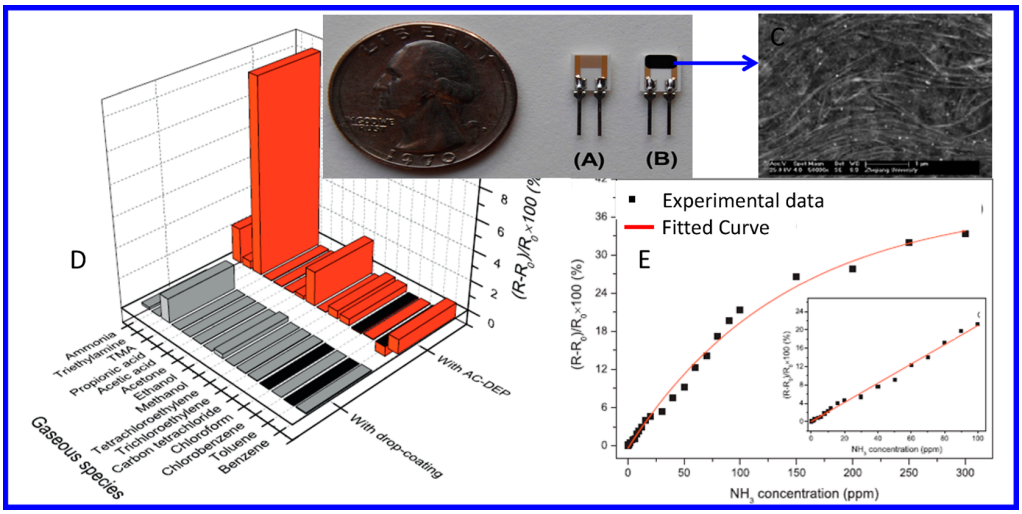


Figure 15. Microelectrodes before (A) and after (B) the deposition of the AC-DEP-assembled PEDOT/PSS-SWCNTs composite film, (C) FE-SEM images of the PEDOT/PSS-SWCNTs composite films doped with SWCNTs (5 mg/mL), (D) selective responses of the drop-coated and AC-DEP-assembled PEDOT/PSSSWCNTs composite films to various vapors at 10 ppm, and (E) the calibration of the sensor response to NH_3 gas at 2−300 ppm and the linear response range (insert). Reprinted with permission from ref 160. Copyright 2013 Elsevier.

thin PANI films, PANI-MWCNT and PANI-SWCNT nanocomposite films, and PANI nanogranules embedded in a polyvinylpyrrolidone matrix fabricated onto an Au integrated microelectrode.[154] These sensing systems were tested for NH_3, H_2, ethanol, methanol, and acetone. It was observed that, at low concentration, PANI thin films (~100 nm thick) showed sensitivity to NH_3 below 0.5 ppm, which is higher than that of nanocomposite films. A nanocomposite thin film of PANI-TiO_2 was fabricated via the self-assembly method by Tai et al.[155] onto a silicon substrate covered with IDE for NH_3 (23−141 ppm) and CO gas sensing applications at 60 °C. The results have shown that the response, reproducibility, and stability of the PANI-TiO_2 thin film to NH_3 is superior than CO gas with a much smaller effect of humidity on the resistance of the PANI/TiO_2 nanocomposite film. Further, sensing parameters of PANI-TiO_2 nanocomposite thin films have revealed faster response (30 s) and recovery rates (40 s), with stabilities of 30 days. This may be due to the presence of TiO_2 nanoparticles, which influence the morphology of the nanocomposite film, resulting in variation of the sensor response-recovery behavior. This group also fabricated hybrid nanocomposites of PANI with TiO_2, SnO_2, and indium oxide (In_2O_3) using the in situ self-assembly technique for NH_3 sensing (23−141 ppm).[156] The results of sensing studies showed that all PANI-based nanocomposite systems had faster response (2−3 s) and recovery (23−50 s) times with better reproducibility (4 cycles) and long-term stability (15 days). The effect of humidity on the response of all these sensors was also investigated. However, PANI-TiO_2 nanocomposite-based sensors showed optimum NH_3 gas-sensing characteristics, due to the compatible energy band gap.

The NH_3 sensors discussed above have limited selectivity and poor detection limits. To fabricate an efficient NH_3 sensor, A PPy-SWCNT hybrid nanocomposite synthesized via chemical polymerization has been used to fabricate a sensitive NH_3 gas sensor.[157] These authors also studied the effect of film thickness, annealing temperature, and SWCNT concentration on the gas sensing characteristics of nanocomposite thin films. Sensors under optimized conditions exhibited a response of 26−276% upon NH_3 exposure ranging from 10 to 800 ppm, with response and recovery times of 22 and 38 s, respectively. Thermal dynamic refluxing and vapor-phase polymerization techniques have been used to encapsulate Pd nanoclusters with PPy to fabricate a nanocomposite film for rapid (14 s) NH_3 sensing.[158] The optical, electrical, and gas sensing properties of a PPy-Pd nanocomposite system have been found to be dependent on the Pd nanoparticle size (15−35 nm), Pd concentration, reaction time, and the morphology of the nanocomposite film. The fabricated sensor exhibited a reversible and reproducible electrical signal with response and recovery times of 14 and 148 s for 1000 ppm. The magnitude of sensor response was 13.9−58.9% for detection from 50 to 2000 ppm at room temperature. The obtained sensing characteristics were better than those of PPy-based NH_3 sensors. The fabricated sensors were also tested for various VOCs and were found to be very selective.

Lim et al.[159] investigated the electrical and NH_3 gas sensing properties of PANI−SWNTs using temperature-dependent resistance and FET transfer characteristics. The sensing response due to the deprotonation of PANI was found to be positive for NH_3 (25−200 ppb) and negative to NO_2 and H_2S. This sensor exhibited sensitivities of 5.8% (NH_3), 1.9% (NO_2), and 3.6% (H_2S) with lower detection limits of 50, 500, and 500 ppb, respectively. However, the poor selectivity of this fabricated sensor limits its further application. Jian et al.[160] reported poly(3,4-ethylenedioxythiophene)/poly-(styrenesulfonate) (PEDOT/PSS) and O_2 plasma-treated SWCNT nanocomposite films as materials of high gas sensing capability at the ppb level. The sensor has been prepared through dielectrophoretic assembly of nanostructured layers of PEDOT/PSS-(SWCNTs) on Au IDEs (Figure 15A,B). The $\pi-\pi$ interactions and electrostatic attraction between PEDOT chains and the SWCNTs result in uniform dispersion of SWCNTs in the PEDOT matrix (Figure 15 C). This gas sensor showed stable and reversible response to NH_3 in the range 2−300 ppm and trimethylamine from 6 to 1000 ppb at room temperature (Figure 15D). The gas sensing properties were

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

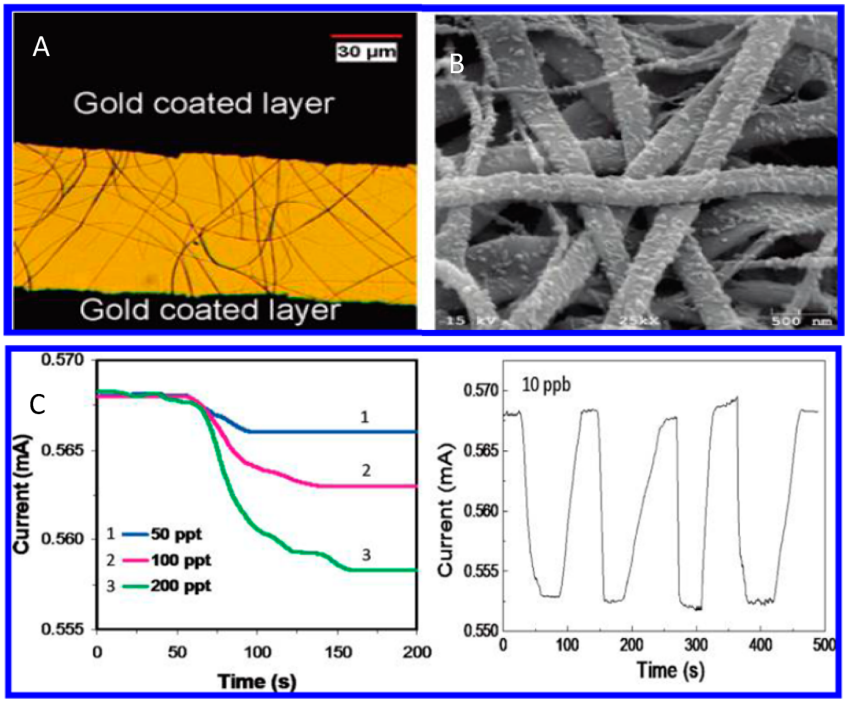


Figure 16. (A) Optical microscope image of the sensor and (B) TiO_2 microfibers encased with PANI nanograins. (C) Current responses of a sensor made of TiO_2 microfibers encased with PANI nanograins to different concentrations of NH_3 gas as a function of time (left) and reproducibility of the sensor exposed to 10 ppb NH_3 gas (right). Reprinted with permission from ref 164. Copyright 2010 American Chemical Society.

found to be dependent on the concentration and alignment of SWCNTs in the PEDOT matrix.

A 2D-ordered, large effective surface area, free-standing and patterned hybrid nanocomposite platform of PANI nanobowl-Au nanoparticles (15 nm) was self-assembled onto polystyrene (PS) spheres at the aqueous/air interface as a template and used for NH_3 detection (0−1600 ppm) by Jiang et al.[161] The sensor with a thickness of ∼100 nm exhibited a fast response time of 5 s with recovery time of 7 s at 100 ppm of NH_3. A one-pot synthesis approach has been used to fabricate Au/CNT-PANI nanocomposites for online monitoring of NH_3 gas.[162] The fabricated Au/CNT-PANI nanocomposite-based sensors showed repeatability upon repeated exposure to NH_3. The improved properties were attributed to fast charge transfer between the electron-donating molecule and CNTs, resulting in a change in conductivity of the system. The sensor exhibited a linear detection range from 200 ppb to 10 ppm, a mean sensitivity of 0.638 (at 25 ppm), a response time of 10 min, and a recovery time of 15 min. Zhang et al.[163] prepared a camphorsulfonic acid (CSA)-doped PANI−SWCNT nanocomposite-based (diameter 17−25 nm) gas sensor using electropolymerization for the selective and sensitive detection of NH_3. The sensor was fabricated using electropolymerization for 2 min. However, the sensor-to-sensor variability was found to be ∼±30% due to batch-to-batch processing. This sensor showed humidity-independent NH_3 sensing responses ranging from 10 to 80 ppb, while the relatively high sensitivity of the sensor to NH_3 concentrations of 400 ppm was demonstrated. The obtained high sensitivity is due to the coordinating roles of the nitrogen atoms of NH_3 gas and the nanocomposite and the deprotonation/protonation process due to the adsorption−desorption of NH_3. The selectivity of the sensor was studied using 1 ppm of NO_2, 3000 ppm of H_2O, and 1 ppm of H_2S. The resistance of SWCNTs increases upon H_2O adsorption, because H_2O molecules donate electrons to hole charge carriers of p-type SWCNTs. Contrary to SWCNTs, the resistance of CSA-doped PANI films decreased upon exposure to H_2O, which was explained by the proton exchange-assisted conduction of electrons mechanism. Thus, the opposite electrical responses of SWCNTs and PANI(CSA) toward H_2O made it possible to eliminate the effect of RH by having just the right balance of PANI(CSA) functionalization. These researchers demonstrated that the effect of cross-interference can be eliminated by accurate control of precursor concentration. Such novel nanoengineered sensing materials are capable of overcoming the shortcomings of conventional sensors.

A highly sensitive (at ppt level) NH_3 sensor was fabricated by Gong et al.[164] P-type conductive PANI nanograins were prepared onto an electrospun n-type semiconductive TiO_2 fiber surface (Figure 16). The structure resulted in two p−n junctions at PANI/TiO_2 interfaces. The forward resistance R1 and bulk resistance R2 are very small compared to the reverse resistance R3. Due to the fact that PANI is much more conductive than the TiO_2 microfiber, the current flows preferably through the conductive PANI particle, when the electric field applied to the reverse bias of the p−n junction is higher than the breakdown voltage. When exposed to NH_3, the bulk resistance of the PANI, R2, increases due to its dedoping, and the current flows solely through the TiO_2 microfibers rather than the PANI nanograins. Therefore, the *I*−*V* response changes from nonohmic to ohmic after exposure to NH_3. The resistance of the p-n heterojunctions combined with the bulk resistance of the PANI nanograins can function as electric current switches when NH_3 gas is absorbed by PANI nanoparticles. As a result, the sensor sensitivity can be

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

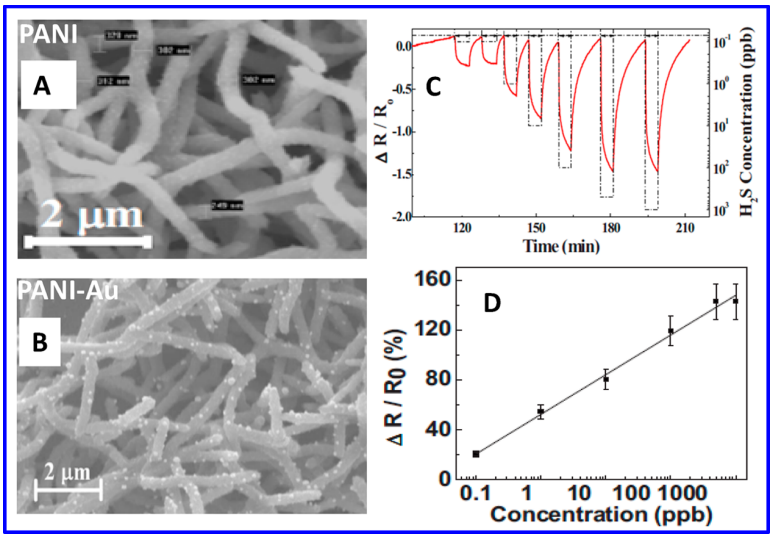


Figure 17. (A) SEM image of a PANI nanowire network grown on Au IDE at the end of second step. (B) SEM image of a gold nanoparticle-functionalized PANI nanowire network across gold microelectrodes. (C) Response and recovery behavior of sensor as a function of 0.1 ppb, 1 ppb, 10 ppb, 100 ppb, 500 ppb, and 1 ppm of H_2S gas. (D) Response R/R_0% of a gold-nanoparticle-functionalized PANI nanowire-network-based chemiresistive sensor as a function of H_2S concentration. Reprinted with permission from ref 168. Copyright 2009 American Institute of Physics.

significantly improved. The detection of 50 ppt of NH_3 gas in air were observed with a sensitivity of 0.004. The repeatability and recovery of the sensor was tested using 10 ppb of NH_3. However, the selectivity of this sensor was not tested, and the observed sensitivity was very low.

4.4. Organic–Inorganic Hybrid Nanocomposites for Hydrogen Disulfide (H_2S) Detection

Hydrogen sulfide (H_2S) is a hazardous toxic gas. At low concentrations, it can lead to personal distress, whereas in a concentration higher than 250 ppm, it may result in loss of consciousness, permanent brain damage, and may cause death. Although a number of techniques have been reported for the estimation of H_2S concentrations, new advanced techniques with high sensitivity, fast response, easy operation, and low cost are still desirable.[165]

Geng et al.[166] prepared PPy/γ-Fe_2O_3 nanocomposites for NH_3 and H_2S gas sensing. It was found that the PPy exhibited maximum sensitivity of 48.5 for NH_3 and 31.4 for H_2S at room temperature. However, a PPy/γ-Fe_2O_3 hybrid nanocomposite annealed at 150 °C showed a maximum sensitivity of 68 for NH_3 and of 406.8 for H_2S, whereas PPy exhibited decreased gas sensitivity at this temperature. The results revealed that PPy/γ-Fe_2O_3 had higher sensitivity and stability than that of PPy, and the working temperature was lower than that of γ-Fe_2O_3. Furthermore, PPy/γ-Fe_2O_3 showed superior properties to those of pure PPy and r-Fe_2O_3. A CSA-doped PANI-CdS nanocomposite prepared using chemical polymerization was used for the selective detection of H_2S (10–100 ppm). This sensor exhibited a maximum response of 76% at 100 ppm and 97.34% stability after 10 days for 40% doping of CSA in the PANI-CdS nanocomposite. The CSA-PANI-CdS sensor exhibited negligible response (2–5%) to NO_2, CH_3OH, C_2H_5OH, and NH_3. Unfortunately, however, this sensor possesses a high recovery time of ~205–413 s.[167]

Due to the fact that PPy and oxide nanoparticles have lower conductivity, Shirsat et al.[168] used metal nanoparticles to decorate synthesized polymer nanowires to achieve higher conductivity and higher sensing surface area. To prepare hybrid networks, PANI nanowires bridging the 3 μm gap between two Au IDEs were synthesized using a two-step galvanostatic electrochemical polymerization technique. These nanowire networks were then functionalized by controlled growth of Au nanoparticles of size ~70–120 nm.[168] The PANI-based sensor exhibited a detection limit of ~500 ppm due to lower conductivity. However, a PANI-Au nanocomposite chemiresistive sensor exhibited an enhanced detection range of 0.1–100 ppb and a detection limit of 0.1 ppb with good selectivity and reproducibility. The improved sensing response has been ascribed to the formation of AuS by the reaction $H_2S + Au \rightarrow AuS + H^{2+}$ and enhanced doping of PANI. The electron transfer from PANI (donor) to Au (acceptor) increases the conduction of PANI (Figure 17).

4.5. Organic–Inorganic Hybrid Nanocomposites for Detection of Nitrogen Oxides (NO_x)

As a result of industrialization, harmful gases such as nitrogen oxides (NO_x), carbon monoxide, and hydrocarbons are of serious concern because of increasing atmospheric pollution and increasing automobile density. The monitoring of these gases is important in order to control their emission from automobiles and industrial plants. NO_x is evolved during the formation of ground-level ozone in the atmosphere, contributes to acid rain, and is an active ingredient in global warming.[52,169] It is also involved in various nitration processes with biomolecules, causing certain forms of cancer and other diseases. Among several oxides of nitrogen present in air, NO and NO_2 are considered the most harmful in terms of both health risk and environmental damage. NO_2 is produced together with NO by oxidation of nitrogenous compounds in fuel combustion.[169] A NO_x sensor that can be attached to the exhaust systems of automobiles could play an important role in monitoring the level of NO_x.[170,171] Therefore, there is an urgent need for efficient sensors to detect NO_x in the ambient environment.

Kaushik et al.[52] and Prakash et al.[172] used WO_3–PANI and hemin/ZnO-PPy nanocomposite thin film sensors, respectively, to detect NO_x gases. The NO_x gas sensing characteristics of the

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

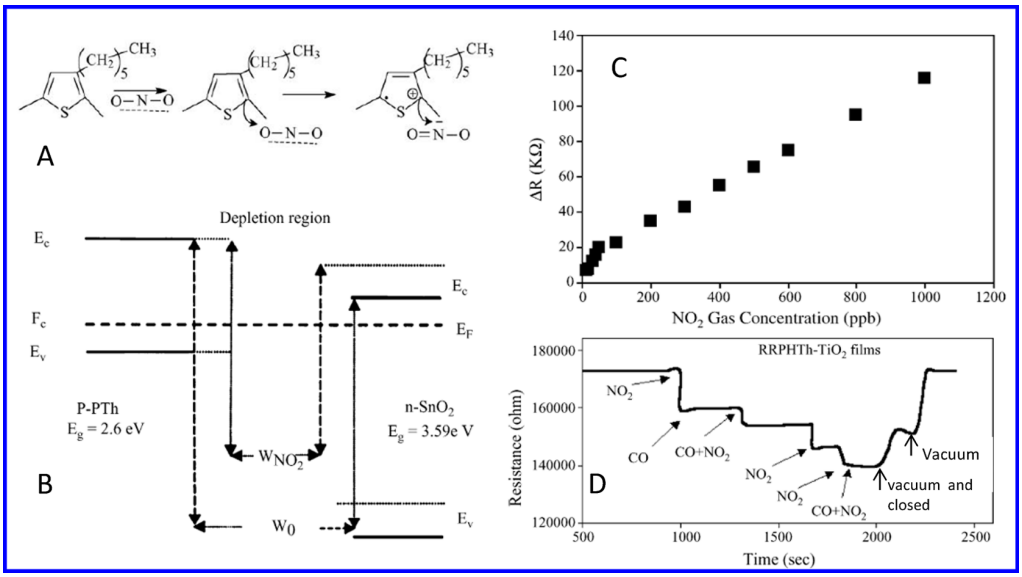


Figure 18. (A) Schematic of the NO_2 reaction with PHTh conducting polymer, (B) energy band of hetero p−n junction with and without NO_2 exposure, (C) resistance change of a RRPHTh−SnO_2 nanocomposite vs NO_2 gas concentration, and (D) resistance change vs time by CO + NO_2 gas concentration in RRPHTh−TiO_2 films. Reprinted with permission from ref 110. Copyright 2005 Elsevier.

sensors were studied by measuring the change in resistance with respect to time. The studies suggested that hemin/ZnO-PPy/Pt electrodes exhibited 3-fold-enhanced electrochemical activity toward detection of NO_x compared to hemin/PPy/Pt electrodes. This sensor exhibited a linear range of 0.8−2000 μM, a sensitivity of 0.04 μM^{-1}, and a detection limit of 0.8 μM at room temperature. Geng et al.[173] reported the application of PPy-ZnO nanocomposites for the sensing of toxic gases such as NH_3, H_2S, and NO_x at different operating temperatures (30, 60, and 90 °C) at gas concentrations of 1000, 1500, and 2000 ppm. PPy-ZnO nanocomposites showed superior selectivity and reversibility to NO_x compared to PPy. Furthermore, the working temperature of the PPy-ZnO composites was much lower than the reported working temperature of ZnO (350−450 °C). PPy-ZnO (10%) showed the maximum sensitivity to NO_x in identical experimental conditions.

The nanocomposites discussed above were prepared by mechanical mixing, which leads to inferior attachment and interaction at the interface. Xu et al.[174] used in situ chemical oxidative polymerization methods to fabricate a SnO_2 hollow sphere (20%)-polythiophene hybrid composite system for NO_2 sensing (10−200 ppm). It was observed that a strong synergetic interaction existed between the SnO_2 hollow spheres and polythiophene, resulting in higher sensitivity with good thermal stability. This gas sensing system showed enhanced sensor response (S = 6), selectivity and a short response time (2 min), but a longer recovery time (~24 h) at a working temperature of 90 °C. These results were attributed to the high surface area of the hybrids and the p-n heterojunction formed between p-type polythiophene and n-type SnO_2 hollow spheres. Sadek et al.[175] used the highly sensitive SAW technique to fabricate sensors for the detection of toxic gases. Nanofibers of PANI/indium oxide (InO_3) composites were deposited onto a layered ZnO/64° YX $LiNbO_3$ SAW transducer. This sensor was tested to detect H_2 (1.1% in 30 s), NO_2 (500 ppm in 24 s), and CO (2.12 ppm in 30 s) gases at 11, 2, and 2.5 kHz, respectively.[175] The observed reproducibility was 1.12% for H_2, 125 ppm for CO, and 510 ppb for NO_2. The sensor exhibited repeatability for H_2 and CO. However, repeatability at higher concentrations of NO_2 is presently an issue due to nanocomposite poisoning.

The deposition of in situ self-assembled LBL (layer-by-layer) films has been adopted by Ram et al.[110] LBL assembly, based on electrostatic or molecular forces, is a unique approach for the formation of molecular architectures by adsorption of consecutively alternating polyelectrolytes. The most substantial advantage of LBL is the accurately controlled layer thickness, wherein the macroscopic properties of the molecular film can be controlled by the size of the microscopic structure of the repeat unit. Ultrathin films of polyhexylthiophene (PHTh), regioregular PHTh (RRPHTh), (PEDOT), PHTh−PEDOT copolymer, sulfonated PANI, PANI−SnO_2, PPy−SnO_2, PEDOT-SnO_2, PHTh−SnO_2, and copolymer (HTh-PEDOT)−SnO_2 were deposited using layer-by-layer self-assembly for NO_2 gas sensing (Figure 18). The results of these studies suggest that the PHTh and its nanocomposite (6 layered films) provide good platforms for NO_2 recognition (10−800 ppb). These nanocomposites were also studied for selectivity by exposing these films to two or more gases at the same time. Among these, the RRPHTh-SnO_2 nanocomposite-based sensor was found to be the most suitable platform for NO_2 gas sensing applications due to its excellent change in conductivity (ΔR changes from 0 to 120 KΩ at 200 ppm) at room temperature.

The sensing response of the nanocomposite can be enhanced by increasing the conductivity, which can be achieved by (i) replacing oxide nanostructures with metallic nanostructures of better conductivity, or (ii) doping of polymer or oxide nanostructures. Fabrication of a PANI/Au/Nafion hybrid nanocomposite film using the CV technique and its application to amperometric NO_2 gas detection was reported by Doet al.[176] NO_2 diffusion into a porous polymer resulted in reduction of mass transfer resistance with increment in cathodic reduction of NO_2. The sensor exhibited a sensitivity of 3.04 μA ppm^{-1} with a switching effect for 0 and 100 ppm concentrations. It was observed that nanocomposites with higher effective surface area and electrochemical activity could be obtained using electro-

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

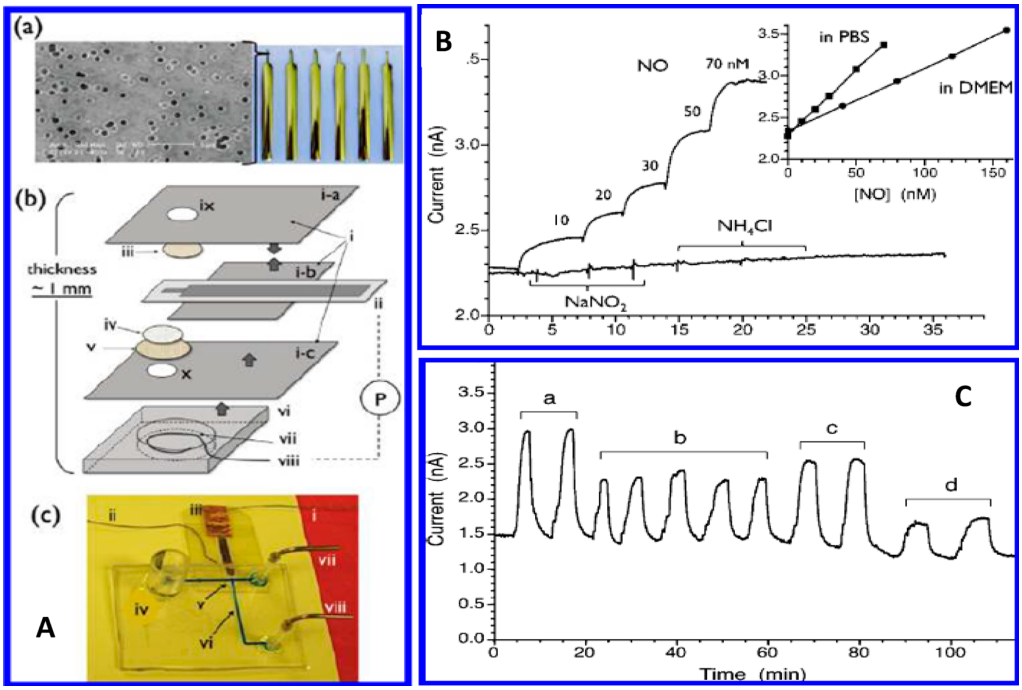


Figure 19. (A) SEM image of electrode and stepwise fabrication of fluidic-based gas sensor, (B) NO sensor response curves exhibiting selectivity over nitrite and ammonia, both species added three times separately; 1, 2, and 3 mM concentrations were obtained after each injection; (inset) NO calibration plots in two different media, and (C) amperometric response of the microfluidic NO detector with flowing test solutions containing different levels of NO; flow rate 50 μL/min, (a) 350 nM, (b) 100 nM, (c) 200 nM, and (d) 40 nM. Reprinted with permission from ref 109. Copyright 2010 American Chemical Society.

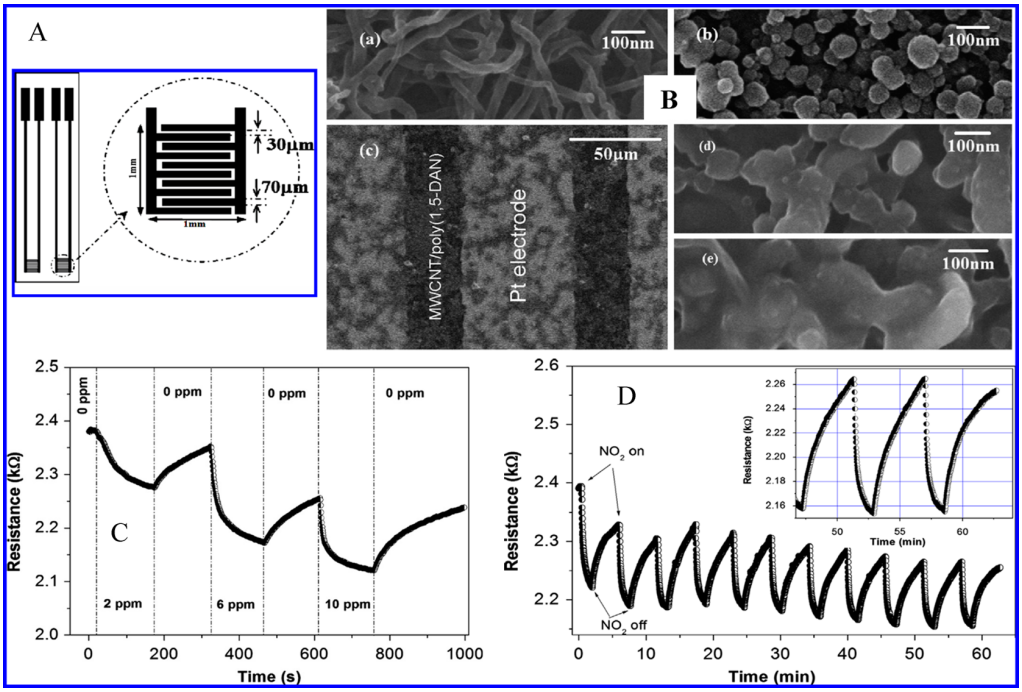


Figure 20. (A) Structure of the IDE electrode, (B) SEM images: MWCNT (a), poly(1,5-DAN) (b), surface of the film IDE/MWCNT/poly(1,5-DAN), and (c) surface morphologies of MWCNT/poly(1,5-DAN) at 10 cycles (d) and at 25 cycles (e). (C) Response of the IDE/MWCNT/poly(1,5-DAN) film formed with 10 polymerization cycles to different concentrations of NO_2. (D) Response−recovery cycles of the IDE/MWCNT/poly(1,5-DAN) film formed with 10 polymerization cycles to 5 ppm of NO_2 and pure air; the inset depicts the last cycles. Reprinted with permission from ref 178. Copyright 2013 Elsevier.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

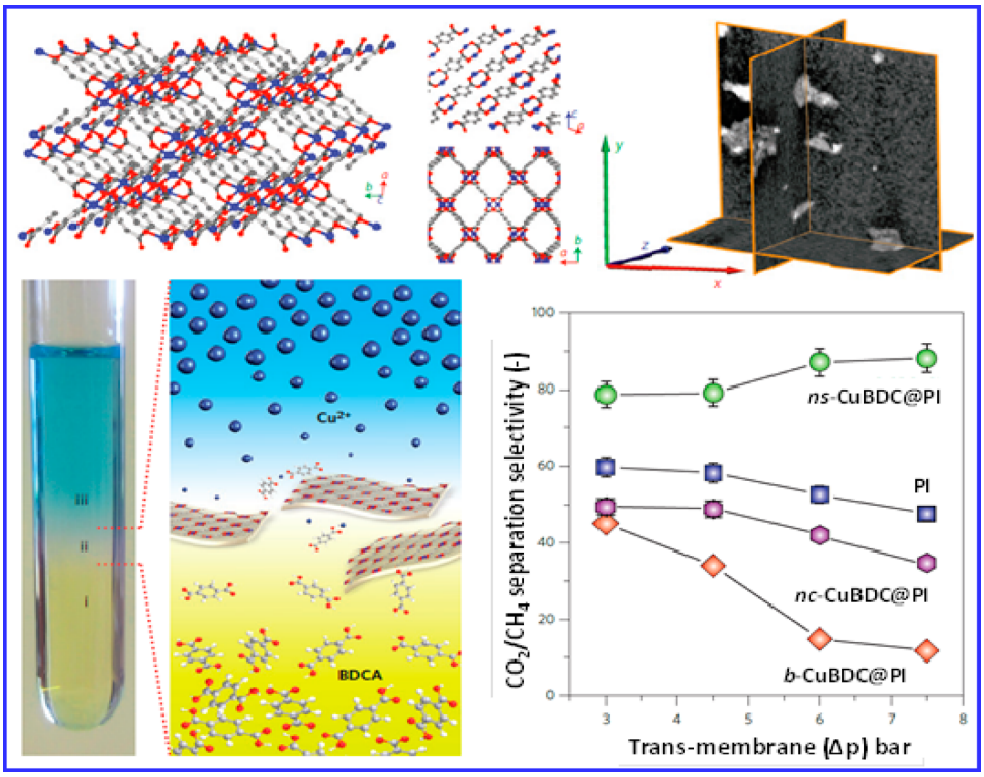


Figure 21. (A) 3D crystalline structure of CuBDC MOF. (B) Spatial arrangement of different liquid layers during the synthesis of CuBDC MOF nanosheets, (i) corresponds to a benzene 1,4-dicarboxylic acid (BDCA) solution, (ii) the solvent spacer layer, and (iii) solution of Cu_2C ions, respectively. (C) 3D-reconstructed FIB–SEM tomogram of a MOF–polymer composite. (D) CO_2 separation performance of a MOF membrane with reference to a standard membrane of polyamide (PI).[183] Reprinted with permission from ref 183. Copyright 2014 Nature Publishing Group.

chemical techniques. Hohercakova et al.[177] prepared Au-PVC nanocomposite electrodes by dispersing Au nanoparticles in a highly plasticized PVC matrix containing a hydrophobic electrolyte, for application to solid-state electrochemical NO_2 sensors. The fabricated Au-PVC nanocomposite electrode, in an array of microelectrodes, exhibited a high catalytic activity. The fabricated sensor showed a linear detection in the 0.6–2.6 ppm (v/v) range with a sensitivity of 400 nA ppm^{-1}, a detection limit of 50 ppb, and a response time of 50 s. The obtained sensing performance was better than that of the Nafion-only NO_2 sensor.

Cha et al.[109] fabricated an amperometric sensor for direct real-time detection of NO produced by cells cultured within the microdevices. The electrochemically deposited Au-hexacyanoferrate layer on the ITO substrate was used to catalyze NO oxidation to nitrite at lower applied potential (0.65–0.75 vs Ag/AgCl) and to obtain a stable current output. The stepwise fabrication of gas sensing architecture is shown in Figure 19A.[109] Furthermore, use of a gas-permeable membrane to separate internal sensor compartments from the sample phase imparts NO selectivity over common interfering agents present in culture media and biological fluids, such as ascorbate, NH_3, etc. (Figure 19B). This reversible amperometric sensor exhibited a linear range of 0–250 nM, a detection limit of ~1 nM (in stirred buffer), a sensitivity of ~10 pA/nM, and stability of 4 weeks. The amperometric response studies as a function of NO concentration were performed using a flow rate of 50 μL min^{-1} at 350, 100, 200, and 40 nM of NO (Figure 19C).[109] This sensor has applications in mechanistic studies and drug testing.

Nguyen et al.[178] prepared an interpenetrated network of MWCNT-poly(1,5-DAN) on IDE for NO_2 sensing (Figure 20). The MWCNT dispersed in Nafion was drop-coated onto Pt-IDE, followed by electro-polymerization of poly(1,5-DAN). Three interpenetrated networks of MWCNT-poly(1,5-DAN) were obtained by operating the electrochemical process voltage between 0.1 and 0.95 V at a scan rate of 50 mV s^{-1} for 5, 10, and 15 cycles. Under NO_2 exposure, the polymerized MWCNT network showed large resistance change while the non-polymerized MWCNT network showed no significant change.

The gas sensing mechanism has been attributed to the removal of unoccupied electrons from the NH_2 groups in the poly(1,5-DAN) backbone by NO_2 gas molecules. The removal of electrons induces the formation of NO_2^- ions and polarons on the polymer which in turn leads to the formation of new double bonds in neighboring positions. The formation of polarons and bipolarons assist charge migration across the backbone, resulting in decreased resistance. It was observed that the sensitivity and selectivity improves with an increase in polymer film thickness, but the response and recovery time is better for films of lower thickness. The observed drawback in this work is that the resistance change did not recover fully in the observed recovery time period. Thus, more efforts are required to obtain improved response-recovery times.

4.6. Organic–Inorganic Hybrid Nanocomposites for Carbon Dioxide (CO_2) Detection

Detection of carbon dioxide (CO_2) concentration has been recognized as very important to preserve a clean environment in an airtight room to control agricultural and biological

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

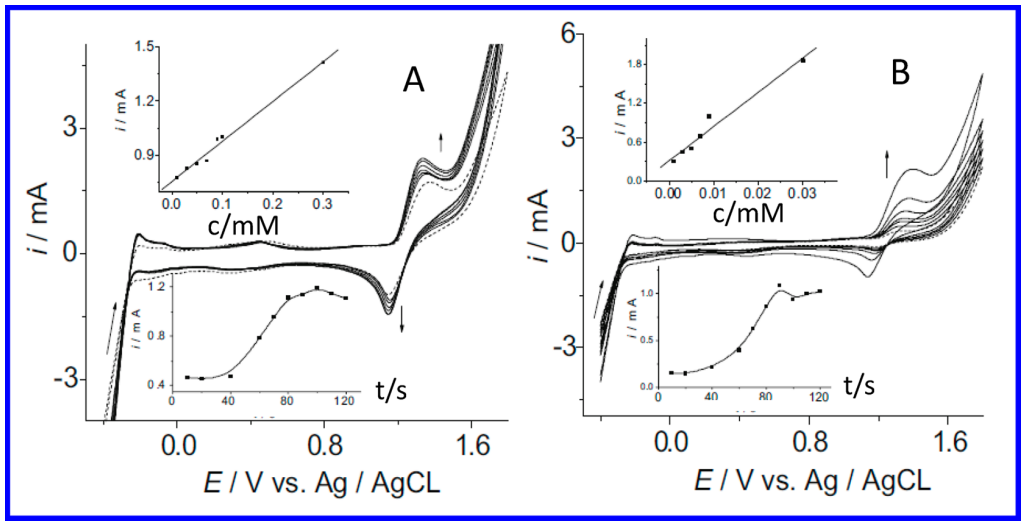


Figure 22. (A) CVs of CO (0.01, 0.03, 0.05, 0.07, 0.09, 0.3, 0.5, and 0.7 mM) at Pt/PAN/MWCNTs/WGE; the upper inset depicts the dependence of reduction peak currents of CO on the concentration. The lower inset shows the accumulation effect of 0.1 mM CO at Pt/PAN/MWCNTs/WGE. (B) CVs of CO (1.0, 3.0, 5.0, 7.0, 9.0, 30.0, and 50.0 μm) at Pt−Ni/PAN/MWCNTs/WGE made in solution concentration ratio of Pt−Ni of 1:1 using constant potential and CV step; the upper inset shows the dependence of oxidation peak currents of CO on the concentration. The lower inset shows the accumulation effect of 0.1 mM CO at Pt−Ni/PAN/MWCNTs/WGE. Buffer: 0.5 M $HClO_4$. Scan: 50 mV s^{-1}. Reprinted with permission from ref 189. Copyright 2009 Elsevier.

processes.[179,180] In this context, infrared spectroscopic and gas chromatographic techniques have been found to be unsuitable because the available apparatus is expensive and large in size. Hence, a number of simplified CO_2 sensors have been proposed that permit the routine measurement of CO_2 gas. The principle of CO_2 sensing depends mainly on the basis of the change in the electrical response of sensing materials as a function of CO_2 concentration; e.g., the electromotive force and capacitance were measured in electrolyte and capacitance type CO_2 sensors, respectively.[181]

Kim et al.[182] studied the gas-sensing properties of organic−inorganic hybrids of poly(amide-6-β-ethylene oxide) (PEBAX) and silica (SiO_2) prepared via in situ polymerization of tetraethoxysilane (TEOS) using sol−gel processes. Gas permeation measurements were conducted at various temperatures with helium, CO_2, O_2, and N_2. The activation energy of CO_2 and its permeation through the hybrid membranes was found to be decreased due to the large contribution of the heat of sorption, while that of N_2 increased resulting from the increase of tortuosity and restricted chain mobility. This was attributed to the strong interaction between CO_2 molecules and SiO_2 domains, which resulted in blocking of additional sorption sites of the polyamide in PEBAX and the organic/inorganic interface. The PEBAX-SiO_2 composite membrane exhibited higher gas permeability coefficients and permselectivity than those of PEBAX, particularly at an elevated temperature: $\alpha CO_2/N_2$ = 72 at 25 °C with PCO_2 = 205 bar and $\alpha CO_2/N_2$ = 44 at 85 °C with PCO_2 = 507 bar for P8S2. There is considerable scope to prepare organic−inorganic hybrid nanocomposites for application to rapid, sensitive, and selective detection of CO_2.

Moreover, organic−inorganic hybrid nanocomposites were used to separate CO_2 from gas mixtures for practical applications.[183,184] A membrane of a metal−organic framework consisting of copper 1,4-benzenedicarboxylate based on a bottom-up approach was prepared to separate CO_2 from CO_2/CH_4 gas mixtures (Figure 21).[183] A free-standing membrane of polysilsesquioxanes (PSSQs), an organic/inorganic hybrid, was also developed for gas separation (He, H, CO_2, O_2, N_2, and CH_4) applications.[184] These membranes were found to show maximum permeability for CO_2 only.

4.7. Organic−Inorganic Hybrid Nanocomposites for Carbon Monoxide (CO) Detection

Carbon monoxide (CO) is a colorless, odorless, and tasteless gas that can inhibit the transportation of oxygen in human blood. CO is one of the most common and dangerous pollutants present in the environment due to its emission from automated vehicles, aircraft, natural gas emission, industrial waste, sewage leaking, mines, etc.[185] Its deleterious effect on human life is well-known.[186] When the level of CO rises beyond the minimum level in the environment, it becomes life-threatening and hazardous.[4] It is thus essential to continuously monitor the CO level in the surroundings and places of work in real time.

The currently available techniques such as gas chromatography, chemical analysis, infrared absorption, electronic noses, and semiconducting oxides, which depend upon sample drawing and analysis using light sources, laser beams, etc. are sophisticated, time-consuming, temperature dependent, and expensive. Recently, several selective, cost-effective, and stable gas sensors using nanocomposites have been reported for CO detection. For CO sensing (7.8 to 1000 ppm), Densakulprasert et al.[187] measured the electrical conductivity of PANI-zeolite nanocomposites as a function of precursor concentration, pore size, and the ion exchange capacity of zeolite. The highest electrical conductivities and sensitivities were obtained with the 13X zeolite, followed by the Y zeolite, and the AlMCM41 zeolite. Zeolite AlMCM41 exhibited poor sensitivity due to large pore size and the lowest Cu^{2+} exchange capacity. Y zeolite and 13X zeolite, however, are known to have similar pore sizes; however, the latter has a greater pore free volume and appropriate location distribution of the Cu^{2+} ions. The observed response time was found to increase with the amount of zeolite in the composites but was inversely related to the

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

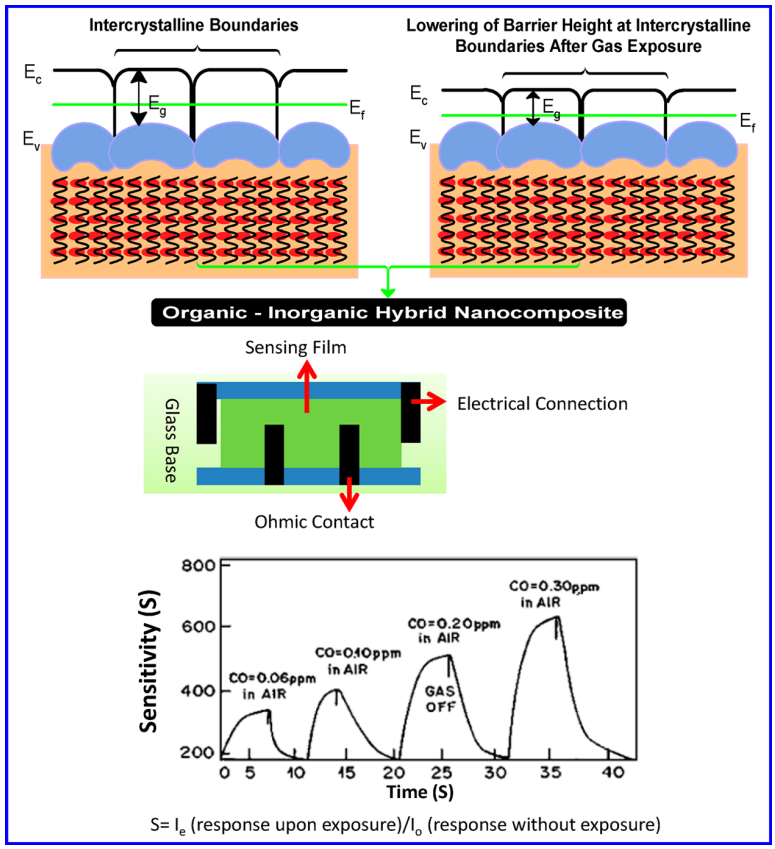


Figure 23. Effect of CO gas on an organic−inorganic hybrid nanocomposite surface. Reprinted with permission from ref 81. Copyright 2004 Elsevier.

amount of ion-exchange capacity. A stable resonance of +C≡O− with the positive charge at the carbon atom withdraws a lone pair of electrons at the amine nitrogen: −NH−. The positive charge at the carbon atom then becomes neutral, and the amine nitrogen then possesses a positive charge and becomes a polaron $-NH^{\bullet+}-$. This amine-to-polaron conversion results in a net increase of the number of positive charge carriers and, therefore, an increase in conductivity.

Ram et al.[188] and Jin et al.[189] have reported improved lower detection limits of 1.7 and 1 ppm for CO, respectively. PANI-MWCNT-based nanocomposites have been used to fabricate an electrochemical gas sensor for CO detection. This sensor works on the basis of bifunctional catalysis and shows enhanced catalytic oxidation and reduction of CO. The observed results revealed detection of CO in the linear range 0.01 to 0.7 mM (Pt/PAN-MWCNTs/WGE) and 10−50 μM (Pt−Ni/PAN-MWCNTs/WGE), respectively, along with a detection limit of 1.7 ppm and 0.1 mM at a scan rate of 50 mV s^{-1}. A response time of 90 s was observed (Figure 22).[189] The ultrathin conducting PANI-SnO_2 and PANI-TiO_2 hybrid nanocomposite films were deposited onto interdigitated electrodes for CO gas (0−1000 ppm) detection, as a function of resistivity change at room temperature. The sensor showed response and recovery times of 20 and 250 s, respectively, at 1 ppm of CO gas.

The rapid and selective detection of CO at ppb level was reported by Mishra et al.[81] using vacuum-deposited PANI-Fe:Al (80:20) nanocomposite thin films. Using these sensors, CO could be detected in the range 0.006 to 0.3 ppm at room temperature. These sensors showed very high sensitivity of the order of 400−600, and response times of 10 s at room temperature. During sensing, CO gas resulted in a reduction of the barrier height at the intercrystallite grain boundary, thus lowering the intercrystallite barrier and increasing the current flow through the sensor. Figure 23 shows the gas sensing behavior of hybrid nanocomposite materials upon exposure to CO gas concentrations. The increase in current output of the sensor was directly proportional to the amount of the exposed gas on the sensor surface. It was observed that vacuum-deposited PANI-Fe:Al thin film sensors were suitable for online detection and continuous monitoring of CO.

5. ORGANIC−INORGANIC HYBRID NANOCOMPOSITES FOR HUMIDITY DETECTION

Next to temperature, humidity is the second most important parameter to monitor in industry. Humidity monitoring is of paramount importance in a number of areas, such as in moisture-sensitive products, storage areas, computer rooms, hospitals, museums, libraries, high-voltage engineering, and accelerator systems.[190,191] Recently, there has been a considerable increase in demand for humidity control in various fields, such as air conditioning systems, electronic devices, the tire industry, the sugar industry, and drying processes for ceramics/food. Moreover, the monitoring of humidity is also important for the fabrication of highly sophisticated integrated circuits (semiconductor industry).[192] A large number of nanostructures ranging from ceramics, semiconductors, and polymer hybrid nanocomposites have been utilized for humidity detection.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Humidity is the most ubiquitous interferent in any realistic environmental sensing situation. Most sensing reports only describe the sensing of gases in the absence of water and likewise, sensing of humidity is done in the absence of any other potentially interfering species. Recently, efforts have been made to prepare gas sensors that will remain unaffected by humidity[193,194] or only selective to humidity, i.e., that remain unaffected by other gases.[195] However, more extensive efforts are required to address this issue. The potential of organic−inorganic hybrid nanocomposites to fabricate humidity sensors is discussed here.

Tandon et al.[196] fabricated conductive nanocomposites of PPy−Fe_3O_4 using an emulsion polymerization method in aqueous solution, for use as gas sensors. This nanocomposite exhibited fairly high sensitivity (980) to relative humidity (80%) and some commonly used gases (N_2, O_2, and CO_2). The sensitivity of PPy-Fe_3O_4 gas sensors increases almost linearly, but the rise is steeper for CO_2 gas, due to the smaller kinetic diameter (0.33 nm) of the CO_2 molecule in comparison to 0.345 nm for O_2 and 0.364 nm for N_2. The fabricated sensor exhibited a response time of 10−15 s and a recovery time of ∼25 s. However, this sensor was not found to be selective. Power et al.[197] fabricated a poly(vinyl alcohol)−Ag nanoparticle (∼22 nm) nanocomposite onto Pt-IDE and utilized it as a sensor to detect humidity selectively. At constant potential, the fabricated nanocomposite exhibited a change in current as a function of %RH from 10% to 60%. The sensor was found to be reversible, selective, and rapid. The authors also proposed the sensor response mechanism, and it correlated well with a second-order Langmuir adsorption model using second-order $S = S/S_s$ ($K_bC/1 + K_bC$), where S_s is relative signal at saturation, K_b is the binding constant, and C is the relative %RH.

Since the chemistry of polymers influences the sensing mechanism, further research focused on tailoring materials with good sensitivity, detection range and detection limits is likely to be very beneficial for device applications. Su et al.[198] described fabrication of a hybrid nanocomposite thin film of PPy-TiO_2 onto an Al substrate using the in situ electrochemical polymerization technique for humidity detection. The sensing behavior and its related mechanism were investigated by measuring the activation energy and by impedance spectroscopy. PPy-TiO_2 hybrid nanocomposite (TiO_2 nanoparticles ∼0.0012 g) thin-film-based sensors (at 15 to 35 °C) exhibited lower hysteresis, sensitivity and good linearity ($Y = -0.306X + 0.84$; $R^2 = 0.94$) with response and recovery times of 20 and 40 s, respectively. The sensing characteristics of this sensor were improved by adding poly[3-(methacrylamino)propyl]-trimethylammonium chloride (PMAPTAC) to a PPy-TiO_2 (0.048 g) nanocomposite using photopolymerization. Enhanced flexibility, sensitivity, linearity ($Y = -0.0665X + 9.57$; $R^2 = 0.997$), hysteresis, a fast response time, a shorter recovery time, and a smaller temperature effect between 15 and 30 °C were obtained.[199] The gas sensing performance of PMAPTAC-PPy-TiO_2 nanocomposite films was not affected by CO and CO_2 gases.

Humidity sensing properties of poly(*o*-anisidine)-WO_3 (POA/WO_3) nanocomposite sensors as a function of resistance versus relative humidity and humidity hysteresis was explored by Patil et al.[200] The humidity sensing properties of the POA were also studied for comparison. Parvatikar et al.[201] prepared WO_3 nanoparticle-based hybrid nanocomposites with PANI for humidity sensing. Patil et al. prepared nanocomposites by mechanical mixing, whereas Parvatikar et al. used the "in situ" deposition technique by placing fine-grade WO_3 nanoparticles in a polymerization mixture of aniline. It was observed that the composite prepared by mechanical mixing of 30 wt % WO_3 in POA showed the maximum response factor of 353 at 85% RH, with narrow hysteresis (6%). The nanocomposites prepared by the in situ method showed better sensing performance. High-temperature conductivity measurements showed thermally activated behavior with a change in resistance with respect to % RH. The almost linear response (10 to 95% RH) of the PANI-WO_3 nanocomposite (50 wt % of WO_3 in PANI) to a broad range of humidity conditions suggested it to be an interesting and competent material for use as a humidity sensor. The superior sensing performance of in situ-prepared nanocomposites over mechanically mixed nanocomposites can be understood in the terms of improved interaction at the interface between inorganic nanoparticles and polymers. The improved interaction at the interface perhaps results in easy and efficient charge transfer and better sensing performance of the nanocomposites.

The impedance spectroscopy technique has been used to investigate humidity sensing and electrical properties of SiO_2 and poly[3-(methacrylamino)propyl]trimethylammonium chloride (poly-MAPTAC)-based nanocomposite sensors.[202] A sensor based on a SiO_2-poly-MAPTAC nanocomposite film showed satisfactory resistance at high humidity atmospheres ∼97% RH with high sensitivity (three orders), better linearity over a wide humidity range (10−90% RH), negligible hysteresis, and response and recovery time (humidification: 60 s, desiccation: 120 s). The activation energy for conduction reduced with water adsorption. Further, different response mechanisms of SiO_2-poly-MAPTAC nanocomposite sensors were suggested at low and high relative humidity. Highly electro-active CNTs were incorporated into polydimethyldiallylammonium chloride (PDDA) to fabricate nanocomposite films between IDEs on a Si/SiO_2 substrate using the layer-by-layer self-assembly technique.[203] The resistance of CNT-PDDA nanocomposite films was found to increase exponentially with an increase in humidity (RH ∼ 20%−98%). In sensing, CNT junctions played an important role in the overall resistance change for water molecule absorption. The sensor exhibited a high sensitivity of 80%, a fast response time of 8 s, and good reversibility (5 cycles). Another impedance-based humidity sensor was reported by Li et al.[204] using nanocomposite thin films of sodium polystyrenesulfonate (NaPSS) and ZnO nanoparticles. The impedance response (as a function of humidity concentration) of the sensor was found to increase by 4 orders of magnitude, over a humidity range of 11−97% RH. The NaPSS-ZnO nanocomposite sensors exhibited sensitivity (2−90% RH), better linearity, low hysteresis (∼2% RH), and a quicker response (2 s for absorption and desorption) than the NaPSS-based sensor. It has been proposed that the improved sensing properties were due to the addition of ZnO nanoparticles to NaPSS.

An optical-fiber-based humidity sensor containing Ag (diameter ∼15 nm)-PANI nanocomposite has been fabricated onto optical fiber clad (length ∼6 mm) using evanescent wave absorption spectroscopy and tested in the range of 5−95% RH.[205] It was observed that reduction of the Ag nanoparticle size leads to improved sensitivity and response time. The fabricated sensor exhibited reversible response with a standard deviation of 1% and response time of ∼30 s with a high recovery time of 90 s. This report focused on the effect of Ag nanoparticle concentration in nanocomposites for sensing

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

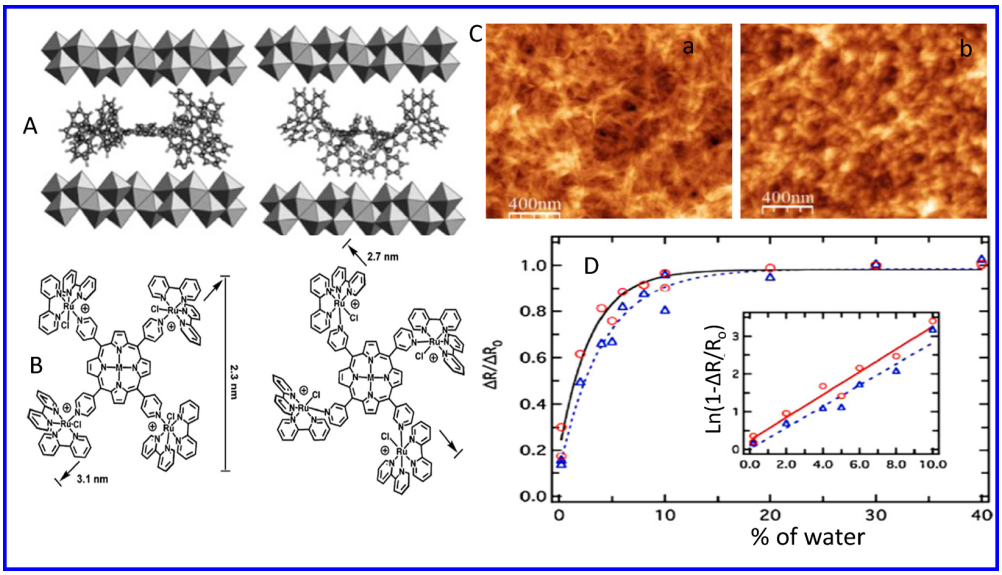


Figure 24. (A) Schematic structures of the Co(III)VXG and Co(IV)VXG host–guest materials. (B) Wireframe structural representation of the cationic Co(IV)TRPyP and Co(III)TRPyP species. The respective chloride salts were used for the preparation of the host–guest nanocomposites with VXG. (C) AFM images of Co(III)VXG (a) and Co(IV)VXG (b) films on mica substrates in a more densely packed region and at the Co(III)VXG film edge. (D) Plot of R/R_0 as a function of the percentage of water (v/v) in ethanol/water mixtures, for IDE Co(III)VXG and Co(IV)VXG sensors. Inset: plot of $\ln(1 - R/R_0)$ vs (% water). Reprinted with permission from ref 207. Copyright 2010 Elsevier.

phenomena. Liu et al.[206] and Timm et al.[207] used IDEs to fabricate sensors of lower dimensions with enhanced sensing performance. Liu et al.[206] deposited MW/SWCNT-PDDA hybrid nanocomposite films onto IDE/Si/SiO_2 substrates by the layer-by-layer self-assembly technique for the sensing of humidity (20–98%). The resistance stability of composite films was found to be improved by thermal annealing. The resistance of composite films increased exponentially with increases in humidity. The sensitivity $[S = (\Delta R/R_{20\%}) \times 100]$ was improved with the increase in the number of layers in the sensor (8 layers). An eight-layered MWCNT-PDDA humidity sensor exhibited a response time of 8 s and a recovery time of 35 s, while in the case of a SWCNT-PDDA sensor, the observed response and recovery times were 8 and 40 s, respectively. A MWCNT-modified sensor showed better sensitivity than SWCNT-based sensors in higher RH ranges (80%). The humidity sensing in CNT-PDDA composite films has been attributed to the junction resistance of CNTs. The sensors were reversible, and temperature was observed to affect the sensitivity of CNT-PDDA composite film humidity sensors.

A nanocomposite film of meso(3/4-pyridyl)porphyrinate Co(III) species coordinated to four $[Ru(bipy)_2Cl]^+$ complexes and vanadium pentoxide (V_2O_5) was deposited onto an Au IDE (Figure 24).[207] This nanocomposite exhibited interesting electrical and electrochemical properties for application to resistive sensors for humidity detection at room temperature. The sensitivity of the fabricated sensor decreased exponentially as a function of water concentration in the ethanol/water system. The sensing phenomenon is based on reversible adsorption/desorption processes involving a defined number of interaction sites, according to the Langmuir equation. These devices were found to be more sensitive toward a lower concentration range (0–10%, v/v) of humidity and allowed quantification of the water content in commercial anhydrous ethanol and hydrated ethanol used as fuel (Figure 24 D). The improved sensing properties have been attributed to the use of IDE in fabrication of sensors. The use of IDE enhances the collection or transfer of charges to the electrode due to the smaller gap between the successive electrodes.

6. CHALLENGES IN ORGANIC–INORGNIC HYBRID NANOCOMPOSITE-BASED GAS SENSING

The applications of organic–inorganic hybrid nanocomposites for environmental monitoring require urgent solutions related to materials processing science and sensor fabrication. Inorganic nanoparticles have a tendency to become agglomerated in solvents, which decreases the effective surface area. This shortcoming can be overcome using different surfactants to modify the surface of nanoparticles for uniform dispersion. Another approach is to use nanostructures with aspect ratios higher than unity, as the agglomeration is reduced in these nanostructures in comparison to spherical nanoparticles. For the fabrication of an efficient sensor, organic–inorganic hybrid nanocomposites need to be deposited or transferred onto an electronically conducting substrate. The challenge is to improve binding of nanocomposites with electrode/conducting platforms and the tuning of energy barrier at the interface for facile charge transport that is required for improved sensing. Depending upon the nature of the electrode and hybrid nanocomposite, various electrode treatments (such as hydrogen/nitrogen/plasma treatment, higher temperature annealing, self-assembled monolayer modification, etc.) have been introduced for efficient sensor platform development. These treatments are likely to result in better attachment of hybrid nanocomposite films to electrodes, along with the lowering of the energy barrier (reduction in the depletion region).

Gas detection at the ppt level requires good analyte adsorption at the nanocomposite sensing surface that can be achieved by developing a better sensing nanomaterial for a given target analyte. The charge carriers generated by the adsorption of the target gas analyte should migrate easily and be collected at the electrode. Due to the low conductivity of

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

Table 1. Gas Sensing Characteristics of Various Hybrid Nanocomposite Systems

sensing platform	sensing characteristics	ref
[NSP] PANI-CNF	[A] food-borne pathogens	208
[R] resistance change	*Bacillus cereus*, *Vibro Parahemolyticus* and *Salmonella*	
[FT] film via drop-coating	[Rm] operate at room temperature and better than metal oxides	
[NSP] PANI-Al diodes	[A] CH_3OH	121
[R] I–V/C-V change	[Rm] CH_4 decreases the reverse bias current and increases barrier height	
[FT] electrochemically		
[NSP] dye-doped silica	[A] CH_3OH	93
[R] fluorescence intensity	[DL] 150 ppm	
[FT] self-assembled sol–gel	[Rm] interaction of gas generates the energy transfer complex and intensity decreases on increasing dose	
[NSP] polysiloxanes-Zn/Ni/Eu	[A] mixtures of VOCs	132
[R] frequency shift	[L] 20–40 ppb	
[FT] spin coated onto quartz crystal	[DL] 0.3 ppb	
	[Rm] good detection range but has sensitivity problems	
[NSP] Cu-PANI	[A] $CHCl_3$	130
[R] resistance change	[L] 10–100 ppm	
[FT] chemical oxidative polymerization	[DL] 10 ppm	
	[S] 1.5–3.5	
	[Rm] resistance changes due to protonation and deprotonation	
[NSP] PVB-Fe_2O_3, PVB-MnO_2	[A] mixtures of VOCs	124
[R] resistance change	[L] 0–5000 ppm	
[FT] chemical oxidative polymerization	[S] sensor of composition 75 Fe_2O_3/25 MnO_2 mol %; shows maximum sensitivity to propanol and toluene	
	[Rm] sensor is not sensitive to CH_3CH_2OH, $CHCl_3$ or C_6H_{14}	
[NSP] PVB-NiO_2/Fe_2O_3	[A] mixtures of VOCs	125
[R] resistance change	[DL] 4000 ppm for $C_6H_5CH_3$ and CH_3CH_2OH	
[FT] self-casting	[Rs] 30 s for CH_3CH_2OH and 40s $C_6H_5CH_3$	
	[Rs] 45 for CH_3CH_2OH and 48 for $C_6H_5CH_3$	
	[Rm] response is crystalline-size-dependent but shows response to many VOCs	
[NSP] Au-polyamide	[A] VOCs	138
[R] optical sensor	[DL] 6000 ppm for CH_3OH, 6000 ppm for CH_3CH_2OH	
[FT] glow discharge vapor deposition	[Rm] Au-polyamide (5×10^{16} Au^+ cm^{-2}) implanted film exhibits interesting optical properties but uses complex technique for response measurement.	
[NSP] PPy-MoO_3	[A] CH_3CH_2OH	139
[R] resistance change	[DL] 0.3%	
[FT] vapor polymerization	[Rs] 2100 s	
	[Rm] resistance of nanocomposite increases on increasing gas dose but operation at high temperature is a major issue	
[NSP] PANI-Ag	[A] CH_3CH_2OH	140
[R] resistance change	[L] 20–225 ppm	
[FT] ultrasound-assisted in situ polymerization then spin-coating	[DL] 100 ppm	
	[Rs] 5–6 min	
	[Rm] response and reaction time is Ag concentration dependent	
[NSP] chitosan-CNTs	[A] CH_3CH_2OH	141
[R] change in conduction	[L] 0–100%	
[FT] layer-by-layer deposition	[DL] high response at 70%	
	[Rm] diffusion of gas leads to tunneling	
[NSP] PDDAC-SnO_2	[A] CH_3CH_2OH	209
[R] change in resistance	[L] 10–200 ppm	
[FT] patterning onto IDEs	[DL] 10 ppm	
	[SL] 60 d	
	[Rs] 88 s	
	[Rm] room-temperature operation, selective and stable sensor; suitable for device fabrication	
[NSP] P3HT/Ng-ZnO	[A] CH_3CH_2OH	210
[R] change in resistance	[L] 50–1000 ppm	
[FT] drop casting	[DL] 50 ppm	
	[Rs] 11 s	
	[Rm] P3HT:0.75 mol % Nb/ZnO sensor exhibited room-temperature sensing; stability and selectivity need to be explored	
[NSP] PANI-Pd	[A] aliphatic alcohols	142
[R] resistance change	[S] 8.9×10^5 Ω/ppm for propanol	
[FT] oxidative polymerization	[DL] 10 ppm	
	[SL] 110 d	

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

Table 1. continued

sensing platform	sensing characteristics	ref
	[Rs] 2 to 12 s	
	[Rm] sensing mechanism depend on size of gas molecule, which affects the barrier height in diffusion	
[NSP] PANI-MoO_3	[A] HCHO and CH_3CHO	134
[R] electrical resistivity	[D] 50 ppm	
[FT] chemical vapor deposition	[T] 30 °C	
	[Rm] PANI used for signal amplification, and gas exposure affects the charge conduction in nanocomposites	
[NSP] PNMA-MoO_3	[A] HCHO and CH_3CHO	135
[R] electrical resistivity	[L] 1−10 ppm	
[FT] chemical vapor deposition	[Rm] response is dependent on interlayer space and detailed analysis of sensing characteristics is lacking	
[NSP] PoANIS-MoO_3	[A] HCHO and CH_3CHO	136
[R] electrical resistivity	[Rm] fabricated sensor has less impurity and exhibits relative sensing performance	
[FT] chemical vapor deposition		
[NSP] PANI-TiO_2	[A] $N(C_2H_5)_3$	131
[R] electrical resistivity	[S] 5.14×10^{-7} mol mL^{-1}	
[FT] spin-coating onto interdigitated electrode	[Rs] 3 min	
	[Rm] protonation- and deprotonation-based sensing, also dependent on TiO_2 concentration; sensing characteristics could be improved	
[NSP] PANI-WO_3	[A] H_2	122
[R] surface acoustic wave	[DL] 7 kHz for 1% H_2	
[FT] sol−gel technique	[L] 0−1%	
	[Rm] sensor is sensitive and reproducible at room temperature over an interval of 5 d; other parameters are not related	
[NSP] PANI-anantase TiO_2	[A] H_2	211
[R] change in resistance	[S] 1.63	
[FT] chemical oxidative polymerization	[L] 0.4−1.0 vol %	
	[DL] 0.8 vol %	
	[Rs] 83 s, [Rv] 130 s	
	[Rm] I−V of sensor exhibited an ohmic characteristic; band gap theory on gap adsorption and desorption is proposed in this work; the stability, selectivity and industrial applications should be explored	
[NSP] PP-AN-ZnO_2	[A] O_2	123
[R] conduction change	[S] 16.6 ppm	
[FT] plasma polymerized	[T] 150 °C	
	[Rs] 3 s	
	[Rm] UV exposure generates photocurrents resulting in increased free electron−hole pair that reduces the electron depletion region	
[NSP] polymer-Sb-SnO_2	[A] HCHO	133
[R] conduction change	[L] 10−50 ppm	
[FT] template sol−gel	[T] 136 °C	
	[DL] 10 ppm	
	[Rm] shows better sensing parameters	
[NSP] P3HT-ZnO-based OTFT	[A] HCHO	212
[R] change in resistance	[L] 10−150 ppm	
[FT] spray deposition	[DL] 4 ppm	
	[SL] 15 d	
	[Rs]	
	[Rm] OTFT-based room-temperature HCHO sensor; sensitivity, response time, and sensitivity was not explained	
[NSP] PANI-CNTs	[A] HCHO, $CHCl_3$, CH_3OH, $C_6H_5CH_3$, H_2O	137
[R] chemoresistive	with flow rate 100 cm^3 min^{-1} to 50 ppm	
[FT] oxidative dispersive polymerization	[Rm] effect of precursor concentration of VOC sensor has been discussed but this research lacks selective detection of VOCs	
[NSP] Pd/Ni-poly(EDOT-*co*-TAA)	[A] CH_3COCH_3 and $C_6H_5CH_3$	126
[R] resistance change	[L] 0−1800 ppm	
[FT] wet chemical route	[DL] 50 ppm for CH_3COCH_3 in case of Pd based sensor	
	[DL] 10 ppm for $C_6H_5CH_3$ in case of Ni based sensor	
	[Rs] 60−70 s for CH_3COCH_3	
	[Rs] 200−300 s for $C_6H_5CH_3$	
	[Rm] fabricated sensor is thickness-dependent (optimized thickness ∼50 nm)	
[NSP] PANI-SnO_2	[A] C_6H_6 and $C_6H_5CH_3$	213
[R] thermogravimetric	PANI-SnO_2 (40%) nanocomposite exhibited maximum response to C_6H_6 (16.9%) and $C_6H_5CH_3$ (19.5%)	
[FT] in situ chemical oxidative polymerization	[Rm] presented research work is only a proof of the concept	
[NSP] PANI-SnO_2	[A] C_6H_6 and $C_6H_5CH_3$	214

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

sensing platform	sensing characteristics	ref
[R] change in voltage	PANI-SnO_2 (40 wt %)-based sensor (10% sensor efficiency) was found more sensitive for C_6H_6	
[FT] physical grinding	[Rm] sensing performance was not explored.	
[NSP] PANI-MWCNTs	[A] hydrocarbons	120
[R] resistance change	[DL] ppm level and maximum sensitivity at 1000 ppm	
[FT] electrochemical polymerization	[Rm] polarity of the aromatic hydrocarbons was found to be responsible for the response; however, sensing parameters could be improved	
[NSP] MWCNTs-PMMA	[A] chloroform, methanol, toluene, and water	144
[R] resistance change	[Rm] sensor is selective for methanol; CNT network in PMMA resulted in a high surface area, increasing conducting network sensitivity by ∼240%	
[FT] layer-by-layer deposition		
[NSP] CdSe-ZnSe and PS−PSMA	[A] THF and DMF	128
[R] resistance change	[L] 0−1000 ppm	
[FT] electrospun technique	[Rs] ∼1 s	
	[SL] 6 months	
	[Rm] sensor found to be sensitive to chloroform and THF	
[NSP] PPy-Ag-GO	[A] H_2O_2	215
[R] amperometric	[S] 0.085 (S/N of 3)	
[FT] electrochemical polymerization	[L] 0.1−5 mM	
	[DL] 1.009 mM	
	[SL] 21 d	
	[Rs]	
	[Rm] sensor is nonenzymatic and found to be selective	
[NSP] Fe:Al-PANI	[A] HCl	82
[R] resistance change	[L] 0.1−80 ppm	
[FT] vacuum evaporation deposition	[Rs] ∼10 s	
	[S] 400−8000	
	[Rm] exploration of vacuum-evaporation-deposited nanostructured films for gas sensor applications; sensing mechanism was explained on the basis of the energy band gap theory	
[NSP] CNF-PPy	[A] NH_3 and HCl	148
[R] resistance change	[L] 10−50 ppm	
[FT] vapor deposition polymerization	[DL] 10 ppm for NH_3 and 20 ppm for HCl	
	[Rm] sensor is reproducible and reversible; sensing characteristics are thickness-dependent and based on protonation and deprotonation	
[NSP] TiO_2-TMPyP	[A] HCl	149
[R] resistance change	[T] 80 °C	
[FT] glancing-angle physical deposition technique	[DL] 0.1 ppm	
	[Rs] 51 s	
	[Rv] 300 s	
	[DL] 10 ppm for NH_3 and 20 ppm for HCl	
	[Rm] sensor exhibits enhanced sensing characteristics	
[NSP] poly(*p*-xylene)-Pb	[A] NH_3	216
[R] resistance change	[Rm] current change was measured at 13 ppm with time	
[FT] vacuum deposition technique		
[NSP] PPy-SWCNTs	[A] NH_3	157
[R] electrical resistance	[L] 10−800 ppm	
[FT] chemical polymerization	[Rs] 22 s	
	[Rv] 38 s	
	[T] 25 °C	
	[Rm] fabrication parameters were optimized for 150 ppm of NH_3	
[NSP] PPy-Pd	[A] NH_3	158
[R] electrical resistance	[L] 50−2000 ppm	
[FT] vapor phase polymerization	[Rs] 14 s	
	[Rv] 148 s for 1000 ppm	
	[Rm] response magnitude ∼13.9−58.9; sensor is selective and response is dependent on size and concentration of Pd	
[NSP] PANI-Au	[A] NH_3	161
[R] electrical resistance	[L] 0−1600 ppm	
[FT] self-assembled onto polystyrene sheet	[Rs] 5 s	
	[Rv] 7 s for 100 ppm	
	[Rm] response results were found to be enhanced	
[NSP] PANI-CNT	[A] NH_3	159
[R] electrical resistance	[DL] 50 ppb	

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

sensing platform	sensing characteristics	ref
[FT] self-assembled onto polystyrene sheet	[L] 0−100 ppm	
	[Rm] sensing mechanism is based on protonation and deprotonation and obtained response is improved but stability is an issue here	
[NSP] Au/CNT-PANI	[A] NH_3	162
[R] electrical resistance	[DL] 200 ppb	
[FT] in situ polymerization	[L] 200 ppb to 10 ppm	
	[SL] a few days	
	[Rm] sensor is reproducible and sensitive	
[NSP] PANI-TiO_2	[A] NH_3	156
[R] electrical resistance	[DL] 23 ppm	
[FT] self-assembled onto interdigited electrode of Si	[L] 23−114 ppm	
	[SL] 30 days	
	[Rm] TiO_2 nanoparticles result in improved electrical and morphological properties in nanocomposites resulting in enhanced sensing	
[NSP] PANI-TiO_2	[A] NH_3	164
[R] electrical resistance	[DL] 50 ppt	
[FT] layer-by-layer self-assembly	[L] 23−114 ppm	
	[SL] 30 days	
	[Rm] TiO_2 nanoparticles improve the electrical (via affecting conduction) and morphological properties in nanocomposites, resulting in enhanced sensing	
[NSP] PANI-TiO_2/SnO_2/In_2O_3	[A] NH_3	156
[R] electrical resistance	[DL] 23 ppm	
[FT] in situ self-assembled	[L] 23−114 ppm	
	[SL] 30 days	
	[Rm] PANI-TiO_2 nanocomposite found to be suitable for NH_3 sensing	
[NSP] PANI-MWCNT/SWCNT	[A] NH_3	154
[R] electrical resistance	[DL] 0.5 ppm	
[FT] in situ self-assembled	[Rm] this report focused on material science prospects and applications in gas sensor development	
[NSP] PANI-ZnO	[A] NH_3	217
[R] change in resistance	[L] 25−100 ppm	
[FT] template approach	[Rm] sensing material preparation was optimized to perform NH_3 sensing	
[NSP] CeO_2@PANI	[A] NH_3	218
[R] change in resistance response	[L] 60.5−50 ppm	
[FT] change in resistance	[DL] 25 ppm	
	[SL] 15 d	
	[Rs] 57 s	
	[Rm] researchers explored related band gap diagram; developed sensor was found to be sensitive, repeatable, and selective	
[NSP] guar gum-Ag	[A] NH_3	219
[R] change in current	[S] 1.5 fold	
[FT] wet chemistry routes	[L] 500 ppt to 1300 ppm	
	[DL] 500 ppt	
	[SL] 90 d	
	[Rs] 50 s	
	[Rm] sensor exhibited lowest detection limit	
[NSP] PPy-Fe_2O_3	[A] H_2S	173
[R] electrical resistance	[Rm] PPy-Fe_2O_3 annealed at 150 °C and found to be sensitive to H_2S	
[FT] in situ self-assembled		
[NSP] PPy-Au	[A] H_2S	168
[R] chemoresistive	[DL] 0.1 ppb	
[FT] electrochemical polymerization	[L] 0.1−100 ppb	
	[Rm] fabricated sensor exhibits superior characteristics to those of PANI-based sensing; however, sensing mechanism is not well explained	
[NSP] PPy-WO_3	[A] H_2S	220
[R] relative change in resistance	[S] 83%	
[FT] in situ polymerization	[L] 100 ppb	
	[DL] 100−1000 ppb	
	[SL] 54 d	
	[Rs] 6 min, [Rv] 210 min	
	[Rm] PPy-WO_3 (0.5 g doping) showed better sensing than PPy and WO_3 based sensors	
[NSP] PETH-WO_3	[A] H_2S	221
[R]	[L] 5−200 ppm	

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

sensing platform	sensing characteristics	ref
[FT] nanosheet of WO_3 synthesized using hydrothermal	[DL] 2 ppm	
nanocomposite synthesized using in situ polymerization	[Rv] 15 s	
	[Rm] developed sensor exhibited low detection limit and extremely rapid recovery time; the sensitivity and shelf life of sensor was not discussed	
[NSP] Cu phthalocyanine-polymer	[A] NO_2	222
[R] adsorption resistive	[Rm] this report explains that the nanocomposite system shows improved properties over that of parent material	
[FT] plasma-activated or laser vacuum technique		
[NSP] PEDT-SnO_2	[A] NO_2	110
[R] resistance change	[L] 10−800 ppb and 100−200 ppm	
[FT] layer-by-layer deposition	[S] more than 2 order	
	[Rm] SnO_2-based composite shows better properties than TiO_2-based sensor due to excellent change in conductivity at room temperature; however, there is scope to improve sensing performance	
[NSP] PANI-Au-Nafion	[A] NO_2	176
[R] amperometric	[L] 0−100 ppm	
[FT] CV technique	[S] 3.04 μA/ppm	
	[Rm] this electrochemical sensor exhibits high BET surface and could be used as an electronic nose	
[NSP] PVC-Au	[A] NO_2	177
[R] solid-state electrochemical	[L] 0.6−2.6 ppm (v/v)	
[FT] self coating	[S] 333 nA/ppm	
	[DL] 50 ppb	
	[Rs] 50 s	
	[SL] 30 days	
	[Rm] this sensor exhibits better properties than Nafion-based sensor for NO_2 detection	
[NSP] polymer-Au/ITO	[A] NO_2	109
[R] amperometric	[DL] 1 nM and working potential is very low	
[FT] self coating	[Rm] sensor can be attached to microfluidic devices to detect pathogens and for drug testing	
[NSP] CLPANI-WO_3	[A] NO_2	52
[R] resistance	[S] 4	
[FT] vacuum evaporation deposition	[Rm] mainly focused on nanocomposite fabrication; this report is lacking in detailed discussion of sensing parameters	
[NSP] CH-*co*-PANI-WO_3	[A] NO_2	223
[R] resistance	[L] 100−500 ppb	
[FT] electrochemical polymerization	[Rs] 8 s	
	[SL] 84 days	
	[Rm] this report lacks description of the related sensing mechanism	
[NSP] PPy-NiO	[A] NO_2	224
[R] resistance response	[L] 10−100 ppm	
[FT] spin coating	[DL] 10 ppm	
	[SL] 30 d	
	[Rs] 129−49 s at 100 ppm	
	[Rm] PPy-NiO (50%) exhibited highest sensitivity	
[NSP] PPy-αFe_2O_3	[A] NO_2	225
[R] resistance change	[S] 56%	
[FT] spin coating	[L] 10−100 ppm	
	[DL] 10 ppm	
	[SL] 40 d	
	[Rs] 89 s, [Rv] 4205 s	
	[Rm] sensing performance was compared with PPy- and WO_3-based NO_2 sensors	
[NSP] PPy-WO_3	[A] NO_2	226
[R] impedance change	[L] 5−100 ppm	
[FT] solid state synthesis	[DL] 5 ppm	
	[SL] 20 d	
	[Rm] response-recovery time and sensitivity of presented sensor is not discussed	
[NSP] PPy-ZnO_2	[A] NO_x	173
[R] resistance	[Rm] PPy (10%)-ZnO shows maximum sensitivity at 1000 ppm.	
[FT] electrochemical polymerization		
[NSP] polythiophene-SnO_2	[A] NO_2	174
[R] resistance	[L] 0−200 ppm	

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

sensing platform	sensing characteristics	ref
[FT] in situ chemical oxidative polymerization	[T] 90 °C [Rm] sensor is very selective, improved properties are due to the high surface area of the hybrid nanocomposite	
[NSP] PEBAX-SiO_2	[A] CO_2	182
[R] resistance	[Rm] this sensor explains the effect of CO_2 on the activation energy on nanocomposite	
[FT] sol−gel technique		
[NSP] PANI-Fe:Al	[A] CO_2	81
[R] resistance	[L] 0.006−0.3 ppm	
[FT] vacuum evaporation technique	[Rs] 10 s [S] 400−600 [Rm] this is a very promising sensor for technological applications, wherein CO_2 reduces the barrier height at the intercrystalline grain boundary	
[NSP] PANI-SiO_2 and PANI-TiO_2	[A] CO_2	188
[R] resistivity change	[L] 0−1000 ppm	
[FT] coating onto interdigitated electrode	[S] 400−600 [Rs] 20 s [Rv] 250 s [Rm] this sensor was found to be suitable for online monitoring of CO_2	
[NSP] PANI-zeolite	[Rm] research explains the effect of pore size on sensing	187
[R] resistivity change		
[FT] heterogeneous mixing		
[NSP] PANI-Pt/Ni	[A] CO_2	189
[R] resistivity change	[L] 0.001−0.7 mM for oxidation peak	
[FT] electrochemical polymerization	[L] 1−50 mM for reduction current [Rm] maximum sensitivity observed at 0.01 mM and the enhanced electrocatalytic property is due to bifunctional catalysis	
[NSP] PPy-Fe_3O_4	[A] H_2O	196
[R] resistivity change	[Rm] this composite shows maximum sensitivity for humidity but also less sensitivity to CO_2, N_2, and O_2; selectivity is a major issue	
[FT] electrochemical polymerization		
[NSP] PANI-TiO_2	[A] H_2O	135
[R] activation energy	[Rs] 20 s	
[FT] electrochemical polymerization	[Rv] 40 s [T] 35 °C [Rm] the sensor shows improved sensitivity	
[NSP] PMAPTAC/PPy-Fe_3O_4	[A] H_2O	136
[R] resistivity change	[Rm] the fabricated sensor shows further enhanced sensitivity and flexibility	
[FT] electrochemical polymerization		
[NSP] NaPSS-ZnO	[A] H_2O	204
[R] impedance	[L] 11−97%	
[FT] electrochemical polymerization	[DL] 2% [Rs] 20 s [Rm] obtained sensing is improved. ZnO facilitates the adsorption and desorption of gas on nanocomposite surface	
[NSP] PANI-Ag	[A] H_2O	205
[R] optical fiber based	[L] 5−95%	
[FT] chemical oxidative polymerization	[Rs] 30 s [Rv] 90 s [Rm] sensor exhibits reversible behavior at 1% humidity	
[NSP] PANI-WO_3	[A] H_2O	201
[R] resistivity	[L] 10−95%	
[FT] heterogeneous mixture pallet	[Rs] 30 s [Rv] 90 s [Rm] humidity increases the conductivity of the nanocomposite due to increased hole concentration; detailed studies of R_s, hysteresis, etc. are lacking	
[NSP] POA−WO_3	[A] humidity	200
[R] resistivity	[SL] 30 d	
[FT] heterogeneous mixture pallet	[Rm] optimized nanocomposite shows % response factor as 353 at 85% humidity; however studies of other parameters are not discussed	
[NSP] polyMAPTAC−SiO_2	[A] H_2O	202
[R] electrical resistance	[L] 10−90%	
[FT] sol−gel technique	[S] the obtained response magnitude is 3 orders higher than polymer [Rm] the activation energy for conduction was found to be reduced when OH is adsorbed onto nanocomposite	
[NSP] PEP−$CoCl_2$	[A] H_2O	227
[R] electrical resistance	[Rm] this sensor was used for real-time analysis and exhibits reversible behavior and longer stability	

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Table 1. continued

sensing platform	sensing characteristics	ref
[FT] spin-coated electrodes		
[NSP] PDDA–CNT	[A] H_2O	203
[R] electrical resistance	[L] 20–98%	
[FT] layer-by-layer self-assembly	[Rm] fabricated sensor lacks detailed analysis of sensing characteristics	
[NSP] $[Ru(bipy)_2Cl]^+$–V_2O_5 xerogel	[A] H_2O	207
[R] electrical resistance	[L] 0–10%	
[FT] spin-coated electrodes	[T] 23 °C	
	[Rm] sensor used for low levels of humidity detection	

organic materials in nanocomposites, fewer charge carriers reach the electrodes. This can perhaps be resolved by doping a suitable dopant (mainly surfactant and metal nanostructures, e.g. graphene, Pt, Ni, Au, Ag, etc.) that may improve the charge migration toward electrodes. Another approach adopted by many researchers is to perhaps deposit the nanocomposite sensing thin film onto IDEs (comb like microstructures). These successive electrodes in IDEs are separated by a spacing of a few microns, which may enhance the collection of charge, resulting in better sensing performance. With fast-advancing engineering and technology, it is now possible to fabricate IDEs with electrode interspacing in the range of tens of nanometers. The application of these IDEs in sensing can enhance the signal significantly and sensing at the ppt level can be achieved.

Organic–inorganic hybrid nanocomposites have bright prospects for application in environmental monitoring, but selectivity is still a major concern. The development of optimized target-analyte-specific hybrid nanocomposites is required. Studies and modifications targeted at the interface between organic host materials and guest inorganic nanostructures hold the key for nanocomposite sensing performance. Thus, the science of interfaces needs more clarity and understanding. Due to abundant humidity in the environment, the effect and interference of humidity on sensing needs to be systematically investigated. Hydrophobic organic conducting materials can be used to overcome the problem of humidity. Table 1 summarizes the sensing characteristics of organic–inorganic hybrid nanocomposite-based sensors for environment monitoring.

7. CONCLUSIONS

This article reviews recent developments in the fabrication of organic–inorganic hybrid nanocomposites based gas sensors for environmental monitoring. These materials have been found to be an excellent alternative to conventional gas sensing materials due to their unusual optical, electrical, and morphological properties that are advantageous for gas sensor development, which can easily be tailored by variation of their synthetic routes or precursors. Among the various hybrid nanocomposite systems, conducting polymers and metal/metal oxide nanocomposite sensors are reviewed in greater detail due to their promising abilities as a platform for environmental gas monitoring. This interesting class of hybrid nanocomposite shows interesting optical, electrical and molecular properties that can be easily and cost-effectively tuned. The review clearly demonstrates that sensing characteristics depend both on the composition of the hybrid materials and the geometrical characteristics of the structures produced. The compositions of the materials, in conjunction with their structural morphology and geometrical characteristics, are under further investigation with an aim to optimize materials and overall performance. The review also highlights the urgent need for making measurements in the presence of humidity to establish the ability of the sensors to function in a real world environment.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: ajeet.npl@gmail.com (A.K.)
*E-mail: rajeshbaboria@gmail.com (R.K.)

Notes

The authors declare no competing financial interest.

Biographies



Ajeet Kaushik received his Ph.D. (chemistry, biosensors; 2010) in collaboration with the National Physical Laboratory and Jamia Milia Islamia, New Delhi, India. Presently, he is a faculty at Centre of Personalized Nanomedicine, Institute of NeuroImmune Pharmacology, Department of Immunology, of Florida International University (FIU), USA. He is exploring smart electro-active magnetic nanocarriers for on-demand site-specific delivery and controlled release of therapeutics across BBB to prevent HIV; a new approach towards neuroAIDS. He has worked as a Post-Doctoral Research Associate in the area of fabrication of BioMEMS-based electrochemical biosensors for biomarker detection using nanostructures at the Department of Electrical and Computer Engineering of FIU. He has also worked at Chungnam National University, Deajeon, Korea as a Visiting Researcher and at Dublin City University, Dublin, Ireland as a Research Assistant. He has published over 70 international research papers jointly in the area of nanostructured platforms for biosensor and drug delivery applications. His main research interest is the fabrication of nanomedicines and wearable sensors for personalized health care.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX



Rajesh Kumar is an Assistant Professor in the Department of Physics at Panjab University, Chandigarh, India. He was born in Gurdaspur district of Punjab, India. He received his Ph.D. from the National Physical Laboratory in collaboration with University of Delhi and the Indian Institute of Technology, Delhi, India in the area of Materials Science (2010). He was awarded the prestigious Raman Fellowship by University Grant Commission, India, in 2013 to undertake research in U.S.A. for 1 year. Presently he is working as a Visiting Assistant Professor in the Department of Electrical and Computer Engineering, Florida International University, Miami, Florida. He has published ca. 20 international papers in the area of synthesis and applications of nanomaterials and organic–inorganic hybrid nanocomposites. His research focus is in interface modification of hybrid nanocomposites for energy-harvesting device development.



Sunil K. Arya received his B.Sc. in chemistry in 2002, MSc. in organic chemistry in 2004, and Ph.D. in chemistry and biosensors in 2008 from the University of Delhi, Delhi, India. He is currently a scientist at the Institute of Microelectronics, A*star, Singapore. He has completed Post-Doctoral research at University of South Florida, Tampa, Florida, and at the National Institute of Nanotechnology and University of Alberta, Edmonton, Alberta, Canada. He has published ca. 45 papers in international journals. His primary interests are in the synthesis and applications of nanomaterials, monolayers, and MEMS for use in gas sensors and electrochemical biosensors. His research focus is in the areas of micro/nanofabrication, materials science, and thin film sensors/biosensors.



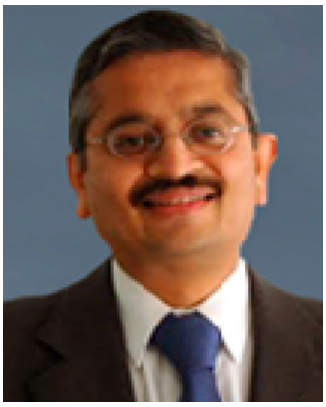
Madhavan Nair received his Ph.D. from Tata Memorial Cancer Center, Bombay University, India, in cancer immunology and trained at the Memorial Sloan Kettering Cancer Center, New York City. He then joined the faculty of the Department of Pediatrics at the University of Michigan, Ann Arbor and subsequently worked at the Department of Medicine and Microbiology at SUNY, Buffalo, New York as a Tenured Professor and Director of Research in Allergy and Immunology. He is a certified Clinical Nutrition Specialist (CNS), Fellow of American College of Nutrition (FACN), and Fellow of the American Academy of Allergy, Asthma and Immunology (FAAAI). Recently, Dr. Nair was named Distinguished Professor and Founding Chair of the Department of Immunology, and Director of the Institute of Neuro-Immune Pharmacology at the FIU College of Medicine, and Associate Dean, Biomedical Research, FIU, Miami, Florida. Nair has published more than 100 research papers, chaired a number of scientific sessions and served in various NIH study sections and committees as chair/member since 1980. His recent research mainly involves the role of different drugs of abuses such as alcohol, morphine, cocaine, and methamphetamine on neuro-AIDS and therapeutic approaches to control neuro-AIDS by specific drugs capable of targeting the brain using nanotechnology.



B. D. Malhotra received his Ph.D. degree from the University of Delhi, Delhi, India in 1980. He has published more than 260 papers in peer-reviewed journals, filed 10 patents, and edited/coedited books on biosensors and polymer electronics. After his stint as Chief Scientist and Head of the DST Centre on Biomolecular Electronics at the National Physical Laboratory, New Delhi, India, he moved to the Delhi Technological University (DTU), India and is currently Professor at the Department of Biotechnology at DTU. He has nearly 30 years of research experience in the field of biomolecular electronics and has guided 24 Ph.D. students to date. His current activities include biosensors, nanobiomaterials, conducting polymers, and ordered molecular assemblies including Langmuir–Blodgett films and self-

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX–XXX

assembled monolayers among others. Dr. Malhotra is a Fellow of the Indian National Science Academy (INSA), the National Academy of Sciences, India (NASI), and an Academician of the Asia-Pacific Academy of Materials (APAM).



Shekhar Bhansali is an Alcatel-Lucent Professor and Chair at the Department of Electrical and Computer Engineering at Florida International University. He received his Ph.D. in Electrical Engineering (1997) from the Royal Melbourne Institute of Technology, Melbourne, Australia. His research interests are in the areas of Bio-MEMS, sensors, nanostructures, and microsystems. He is the recipient of an NSF CAREER Award, an Alfred P. Sloan Foundation Outstanding Mentor Award, FGLSAMP Outstanding Mentor Award, FEF Outstanding Mentor Award, and USF Outstanding Researcher Award. He holds 16 patents in the area of biosensors, nanostructures, and microsystems technologies; has coauthored over 200 journal articles and conference publications; and is the editor of the book "MEMS for Biomedical Applications".

ACKNOWLEDGMENTS

This work was supported by NIH Grants RO1-DA027049, R21-MH101025, RO1-MH085259, and RO1-DA034547 and NSF Grant ERC-ASSIST No. 1160483. DST (DST/TSG/ME/2008/18 and DBT (DBT/GAP070832) is also acknowledged. R.K. acknowledges the UGC, India for the award of Raman Fellowship (ID 1001, 2013-14).

GLOSSARY

$(CH_3)_2CHOH$	isopropanol
[A]	analyte
[DL]	detection Limit
[FT]	fabrication technique
[L]	linear range
[NSP]	nanostructured platform
[R]	response
[Rm]	remarks
[Rs]	response time
[Rv]	recovery time
[S]	sensitivity
[SL]	shelf life
AN	acrylonitrile
ATO	antimony-doped tin oxide
$C_2H_5CH_3$	toluene
C_4H_8O	tetrahydrofuran
C_5H_5N	pyridine
C_6H_{14}	hexane
C_8H_{18}	octane
CdSe	cadmium selenide
CeO_2	cerium oxide
CFC	chlorofluorocarbon
CH	chitosan
$CH_3(CH_2)_3NH_2$	*n*-butylamine
$CH_3CH_2CH_3$	propane
CH_3CH_2OH	ethanol
CH_3CH_2OH	propanol
CH_3CHO	acetaldehyde
CH_3COCH_3	acetone
CH_3OH	methanol
CH_4	methane
$CHCl_3$	chloroform
ClO_4	perchlorate
CLPANI	cross-linked polyaniline
CNF	carbon nanofiber
CNT	carbon nanotube
CO	carbon monoxide
$CoCl_2$	cobalt chloride
COOH	carboxylic acid
CSA	camphorsulfonic acid
CuPc	copper phthalocyanine
DMF	dimethylformamide
EDOT-*co*-TAA	3,4-ethylenedioxythiophene-*co*-thiophene-3-acetic acid
EPA	environment protection agency
Fe_3O_4/Fe_2O_3	iron oxide
FET	field effect transistor
GO	graphene oxide
H_2O_2	hydrogen peroxide
H_2S	hydrogen sulfide
HCHO	formaldehyde
IDEs	interdigitated electrodes
In_2O_3	indium oxide
ITO	indium−tin-oxide
MAcoBMA	methacrylic acid-*co*-*tert*-butyl methacrylate
MAPTAC	[3-(methacrylamino)propyl]-trimethylammonium chloride
MnO_2	manganese oxide
MoO_3	molybdenum oxide
$N(C_2H_5)_3$	triethylamine
NaPSS	sodium polystyrenesulfonate
NBB	nanobuilding block
NH_3	ammonia
NiO	nickel oxide
NIR	near-infrared
NO_x	nitrogen oxides
OTFT	organic thin-film transistor
P3HT	poly(3-hexylthiophene)
PANI	polyaniline
PDDA	polydimethyldiallylammonium
PDDAC	polydiallyldimethylammonium chloride
PEBAX	poly(amide-6-*b*-ethylene oxide)
PEDT	poly(dioxythiophene)
PEHT	polythiophene
PEO	poly(ethylene oxide)
PHTh	polyhexathiophene
PMMA	poly(methyl methacrylate)
PNMA	poly(*N*-methylaniline)
PPV	polyvinyl chloride
PPy	polypyrrole
PSMA	polystyrene-*co*-maleic anhydride
PSSQs	polysilsesquioxanes
PVB	polyvinyl butyral

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

QCM	quartz crystal microbalance
QD	quantum dots
RH	relative humidity
SAW	surface acoustic waves
SnO_2	tin oxide
SPR	surface plasmon resonance
TEOS	tetraethoxysilane
THF	tetrahydrofuran
TiO_2	titanium oxide
TSMRs	thickness-shear mode resonators
UV−vis	ultraviolet visible
V_2O_5	vanadium oxide
VDP	vapor deposition polymerization
VOCs	volatile organic compounds
WO_3	tungsten oxide
ZnO	zinc oxide
ZnS	zinc sulfide
ZrO_2	zirconium oxide

REFERENCES

(1) Zimmerling, R.; Dämmgen, U.; Küsters, A.; Grünhage, L.; Jäger, H.-J. Response of a Grassland Ecosystem to Air Pollutants. IV. The chemical climate: Concentrations of Relevant Non-Criteria Pollutants (Trace Gases and Aerosols). *Environ. Pollut.* **1996**, *91*, 139−147.

(2) Odlyha, M.; Foster, G. M.; Cohen, N. S.; Sitwell, C.; Bullock, L. Microclimate Monitoring of Indoor Environments Using Piezoelectric Quartz Crystal Humidity Sensors. *J. Environ. Monit.* **2000**, *2*, 127−131.

(3) Emil J, B., Jr. Indoor Pollution and its Impact on Respiratory Health. *Ann. Allergy, Asthma, Immunol.* **2001**, *87*, 33−40.

(4) Becker, T.; Mühlberger, S.; Braunmühl, C. B.-V.; Müller, G.; Ziemann, T.; Hechtenberg, K. V. Air Pollution Monitoring Using Tin-Oxide-Based Microreactor Systems. *Sens. Actuators, B* **2000**, *69*, 108−119.

(5) Windmiller, J. R.; Wang, J. Wearable Electrochemical Sensors and Biosensors: A Review. *Electroanalysis* **2013**, *25*, 29−46.

(6) Bogue, R. Emerald Article: Terrorism and Military Actions Pose the Ultimate Challenge to Gas Sensing. *Sens. Rev.* **2011**, *31*, 6−12.

(7) Lim, T. C. *Nanosensors: Theory and Applications in Industry, Healthcare and Defense*. Taylor and Francis: London, 2010.

(8) Korotcenkov, G. *Chemical Sensors: Comprehensive Sensor Technologies*; Momentum Press: New York, 2011; Vol. *6*.

(9) Korotcenkov, G. *Chemical Sensors Comprehensive Sensor Technologies*; Momentum Press: New York, 2011; Vol. *4*.

(10) Janata, J. Introduction: Modern Topics in Chemical Sensing. *Chem. Rev.* **2008**, *108*, 327−328.

(11) Korotcenkov, G. *Chemical Sensors*; Momentum Press: New York, 2010; Vol. *1*.

(12) Stetter, J. R.; Li, J. Amperometric Gas Sensors-A Review. *Chem. Rev.* **2008**, *108*, 352−366.

(13) Blair, R.; Shepherd, H.; Faltens, T.; Haussmann, P. C.; Kaner, R. B.; Tolbert, S. H.; Huang, J.; Virji, S.; Weiller, B. H. Construction of a Polyaniline Nanofiber Gas Sensor. *J. Chem. Educ.* **2008**, *85*, 1102.

(14) Xie, D.; Jiang, Y.; Pan, W.; Li, D.; Wu, Z.; Li, Y. Fabrication and Characterization of Polyaniline Based Gas Sensor by Ultra-Thin Film Technology. *Sens. Actuators, B* **2002**, *81*, 158−164.

(15) Jun, H.-K.; Hoh, Y.-S.; Lee, B.-S.; Lee, S.-T.; Lim, J.-O.; Lee, D.-D.; Huh, J.-S. Electrical Properties of Polypyrrole Gas Sensors Fabricated Under Various Pretreatment Conditions. *Sens. Actuators, B* **2003**, *96*, 576−581.

(16) Chang, J. B.; Liu, V.; Subramanian, V.; Sivula, K.; Luscombe, C.; Murphy, A.; Liu, J.; Frechet, J. M. Printable Polythiophene Gas Sensor Array for Low-Cost Electronic Noses. *J. Appl. Phys.* **2006**, *100*, 014506.

(17) Saxena, V.; Malhotra, B. D. Prospects of Conducting Polymers in Molecular Electronics. *Curr. Appl. Phys.* **2003**, *3*, 293−305.

(18) MacDiarmid, A. G. "Synthetic Metals": A Novel Role for Organic Polymers (Nobel Lecture). *Angew. Chem., Int. Ed.* **2001**, *40*, 2581−2590.

(19) Liu, T.; Burger, C.; Chu, B. Nanofabrication in Polymer Matrices. *Prog. Polym. Sci.* **2003**, *28*, 5−26.

(20) Inzelt, G. *Applications of Conducting Polymers Conducting Polymers*. Springer: Berlin, 2012; pp 245−293.

(21) Korotcenkov, G. *Chemical Sensors Fundamentals of Sensing Materials*; Momentum Press: New York, 2011; Vol. *3*.

(22) Adhikari, B.; Majumdar, S. Polymers in Sensor Applications. *Prog. Polym. Sci.* **2004**, *29*, 699−766.

(23) Barisci, J. N.; Wallace, G. G.; Andrews, M. K.; Partridge, A. C.; Harris, P. D. Conducting Polymer Sensors for Monitoring Aromatic Hydrocarbons Using an Electronic Nose. *Sens. Actuators, B* **2002**, *84*, 252−257.

(24) Bai, H.; Shi, G. Gas Sensors Based on Conducting Polymers. *Sensors* **2007**, *7*, 267−307.

(25) Haynes, A.; Gouma, P.-I. Polyaniline-Based Environmental Gas Sensors for Environment, *Health and Security*; Baraton, M.-I., Ed.; Springer: Netherlands, 2009; pp 451−459.

(26) Janata, J.; Josowicz, M. Conducting Polymers in Electronic Chemical Sensors. *Nat. Mater.* **2003**, *2*, 19−24.

(27) Brailsford, A. D.; Yussouff, M.; Logothetis, E. M. A First Principles Model of Metal Oxide Gas Sensors for Measuring Combustibles. *Sens. Actuators, B* **1998**, *49*, 93−100.

(28) Capone, S.; Siciliano, P.; Quaranta, F.; Rella, R.; Epifani, M.; Vasanelli, L. Analysis of Vapours And Foods By Means of an Electronic Nose Based on A Sol−Gel Metal Oxide Sensors Array. *Sens. Actuators, B* **2000**, *69*, 230−235.

(29) Ehrmann, S.; Jüngst, J.; Goschnick, J.; Everhard, D. Application of a Gas Sensor Microarray to Human Breath Analysis. *Sens. Actuators, B* **2000**, *65*, 247−249.

(30) Tsujita, W.; Yoshino, A.; Ishida, H.; Moriizumi, T. Gas Sensor Network for Air-Pollution Monitoring. *Sens. Actuators, B* **2005**, *110*, 304−311.

(31) Barsan, N.; Weimar, U. Conduction Model of Metal Oxide Gas Sensors. *J. Electroceram.* **2001**, *7*, 143−167.

(32) Barsan, N.; Koziej, D.; Weimar, U. Metal Oxide-Based Gas Sensor Research: How to? *Sens. Actuators, B* **2007**, *121*, 18−35.

(33) Tan, O. K.; Cao, W.; Hu, Y.; Zhu, W. Nano-Structured Oxide Semiconductor Materials for Gas-Sensing Applications. *Ceram. Int.* **2004**, *30*, 1127−1133.

(34) Tan, O. K.; Cao, W.; Hu, Y.; Zhu, W. Nanostructured Oxides by High-Energy Ball Milling Technique: Application as Gas Sensing Materials. *Solid State Ionics* **2004**, *172*, 309−316.

(35) In, M.; Gérardin, C.; Lambard, J.; Sanchez, C. Transition Metal Based Hybrid Organic-Inorganic Copolymers. *J. Sol-Gel Sci. Technol.* **1995**, *5*, 101−114.

(36) Judeinstein, P.; Sanchez, C. Hybrid Organic−Inorganic Materials: A Land of Multidisciplinarity. *J. Mater. Chem.* **1996**, *6*, 511−525.

(37) Lebeau, B.; Sanchez, C. Sol-Gel Derived Hybrid Inorganic-Organic Nanocomposites for Optics. *Curr. Opin. Solid State Mater. Sci.* **1999**, *4*, 11−23.

(38) Sanchez, C.; Soler-Illia, G. d. A.; Ribot, F.; Lalot, T.; Mayer, C.; Cabuil, V. Designed Hybrid Organic-Inorganic Nanocomposites from Functional Nanobuilding Blocks. *Chem. Mater.* **2001**, *13*, 3061−3083.

(39) Sanchez, C.; Lebeau, B.; Chaput, F.; Boilot, J. P. Optical Properties of Functional Hybrid Organic−Inorganic Nanocomposites. *Adv. Mater.* **2003**, *15*, 1969−1994.

(40) Sanchez, C.; Soler-Illia, G. J.; Ribot, F.; Grosso, D. Design of Functional Nano-Structured Materials Through the use of Controlled Hybrid Organic−Inorganic Interfaces. *C. R. Chim.* **2003**, *6*, 1131−1151.

(41) Mammeri, F.; Le Bourhis, E.; Rozes, L.; Sanchez, C. Mechanical Properties of Hybrid Organic−Inorganic Materials. *J. Mater. Chem.* **2005**, *15*, 3787−3811.

(42) Nicole, L.; Boissière, C.; Grosso, D.; Quach, A.; Sanchez, C. Mesostructured Hybrid Organic−Inorganic Thin Films. *J. Mater. Chem.* **2005**, *15*, 3598−3627.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

(43) Sanchez, C.; Julián, B.; Belleville, P.; Popall, M. Applications of Hybrid Organic−Inorganic Nanocomposites. *J. Mater. Chem.* **2005**, *15*, 3559−3592.
(44) Grosso, D.; Ribot, F.; Boissiere, C.; Sanchez, C. Molecular and Supramolecular Dynamics of Hybrid Organic−Inorganic Interfaces for The Rational Construction of Advanced Hybrid Nanomaterials. *Chem. Soc. Rev.* **2011**, *40*, 829−848.
(45) Sanchez, C.; Belleville, P.; Popall, M.; Nicole, L. Applications of Advanced Hybrid Organic−Inorganic Nanomaterials: From Laboratory to Market. *Chem. Soc. Rev.* **2011**, *40*, 696−753.
(46) Sanchez, C.; Shea, K. J.; Kitagawa, S. Recent Progress in Hybrid Materials Science. *Chem. Soc. Rev.* **2011**, *40*, 471−472.
(47) Sanchez, C.; Boissiere, C.; Cassaignon, S.; Chaneac, C.; Durupthy, O.; Faustini, M.; Grosso, D.; Laberty-Robert, C.; Nicole, L.; Portehault, D. Molecular Engineering of Functional Inorganic and Hybrid Materials. *Chem. Mater.* **2013**, *26*, 221−238.
(48) Sanchez, C.; Lebeau, B. Design and properties of Hybrid Organic−Inorganic Nanocomposites for Photonics. *MRS Bull.* **2001**, *26*, 377−387.
(49) Lebeau, B.; Innocenzi, P. Hybrid Materials for Optics and Photonics. *Chem. Soc. Rev.* **2011**, *40*, 886−906.
(50) Tan, J. C.; Cheetham, A. K. Mechanical Properties of Hybrid Inorganic-Organic Framework Materials: Establishing Fundamental Structure-Property Relationships. *Chem. Soc. Rev.* **2011**, *40*, 1059−1080.
(51) Sanchez, C.; Boissière, C.; Cassaignon, S.; Chaneac, C.; Durupthy, O.; Faustini, M.; Grosso, D.; Laberty-Robert, C.; Nicole, L.; Portehault, D. Molecular Engineering of Functional Inorganic and Hybrid Materials. *Chem. Mater.* **2013**, *26*, 221−238.
(52) Kaushik, A.; Khan, R.; Gupta, V.; Malhotra, B. D.; Ahmad, S.; Singh, S. P. Hybrid Cross-Linked Polyaniline-WO_3 Nanocomposite Thin Film for NOx Gas Sensing. *J. Nanosci. Nanotechnol.* **2009**, *9*, 1792−1796.
(53) Nastase, C.; Nastase, F.; Vaseashta, A.; Stamatin, I. Nanocomposites based on Functionalized Nanotubes in Polyaniline Matrix by Plasma Polymerization. *Prog. Solid State Chem.* **2006**, *34*, 181−189.
(54) Gómez-Romero, P.; Chojak, M.; Cuentas-Gallegos, K.; Asensio, J. A.; Kulesza, P. J.; Casañ-Pastor, N.; Lira-Cantú, M. Hybrid Organic−Inorganic Nanocomposite Materials For Application in Solid State Electrochemical Supercapacitors. *Electrochem. Commun.* **2003**, *5*, 149−153.
(55) Shi, J.-l.; Hua, Z.-l.; Zhang, L.-x. Nanocomposites from Ordered Mesoporous Materials. *J. Mater. Chem.* **2004**, *14*, 795−806.
(56) Shchipunov, Y. A.; Karpenko, T. Y. Y.; Krekoten, A. V. Hybrid Organic-Inorganic Nanocomposites Fabricated With a Novel Biocompatible Precursor using Sol-Gel Processing. *Compos. Interfaces* **2005**, *11*, 587−607.
(57) Carotenuto, G.; Nicolais, L.; Martorana, B.; Perlo, P. *Metal-Polymer Nanocomposite Synthesis: Novel ex-situ and in-situ Approaches in Metal−Polymer Nanocomposites*; John Wiley & Sons, Inc.: New York, 2005; pp 155−181.
(58) Kaushik, A.; Kumar, J.; Tiwari, M. K.; Khan, R.; Malhotra, B. D.; Gupta, V.; Singh, S. P. Fabrication and Characterization of Polyaniline-ZnO Hybrid Nanocomposite Thin Films. *J. Nanosci. Nanotechnol.* **2008**, *8*, 1757−1761.
(59) Burghard, Z.; Tucic, A.; Jeurgens, L. P. H.; Hoffmann, R. C.; Bill, J.; Aldinger, F. Nanomechanical Properties of Bioinspired Organic−Inorganic Composite Films. *Adv. Mater.* **2007**, *19*, 970−974.
(60) Godovsky, D. Device Applications of Polymer-Nanocomposites Biopolymers·PVA Hydrogels. *Anionic Polymerisation Nanocomposites*; Springer: Berlin, 2000; Vol. *153*, pp 163−205.
(61) Hatchett, D. W.; Josowicz, M. Composites of Intrinsically Conducting Polymers as Sensing Nanomaterials. *Chem. Rev.* **2008**, *108*, 746−769.
(62) Guido, K. Concepts for the Incorporation of Inorganic Building Blocks into Organic Polymers on a Nanoscale. *Prog. Polym. Sci.* **2003**, *28*, 83−114.
(63) Mitzi, D. B.; Chondroudis, K.; Kagan, C. R. Organic-Inorganic Electronics. *IBM J. Res. Dev.* **2001**, *45*, 29−45.
(64) Mitzi, D. B. Thin-Film Deposition of Organic−Inorganic Hybrid Materials. *Chem. Mater.* **2001**, *13*, 3283−3298.
(65) Sukpirom, N.; Lerner, M. M. Preparation of Organic−Inorganic Nanocomposites with a Layered Titanate. *Chem. Mater.* **2001**, *13*, 2179−2185.
(66) Portney, N. G.; Singh, K.; Chaudhary, S.; Destito, G.; Schneemann, A.; Manchester, M.; Ozkan, M. Organic and Inorganic Nanoparticle Hybrids. *Langmuir* **2005**, *21*, 2098−2103.
(67) Vivekchand, S. R. C.; Kam, K. C.; Gundiah, G.; Govindaraj, A.; Cheetham, A. K.; Rao, C. N. R. Electrical Properties of Inorganic Nanowire-Polymer Composites. *J. Mater. Chem.* **2005**, *15*, 4922−4927.
(68) Kruk, M.; Jaroniec, M. Gas Adsorption Characterization of Ordered Organic−Inorganic Nanocomposite Materials. *Chem. Mater.* **2001**, *13*, 3169−3183.
(69) Tran-Thi, T.-H.; Dagnelie, R.; Crunaire, S.; Nicole, L. Optical Chemical Sensors Based on Hybrid Organic-Inorganic Sol-Gel Nanoreactors. *Chem. Soc. Rev.* **2011**, *40*, 621−639.
(70) Ferey, G.; Serre, C.; Devic, T.; Maurin, G.; Jobic, H.; Llewellyn, P. L.; De Weireld, G.; Vimont, A.; Daturi, M.; Chang, J.-S. Why Hybrid Porous Solids Capture Greenhouse Gases? Chem. *Soc. Rev.* **2011**, *40*, 550−562.
(71) Han, J. T.; Zheng, Y.; Cho, J. H.; Xu, X.; Cho, K. Stable Superhydrophobic Organic−Inorganic Hybrid Films by Electrostatic Self-Assembly. *J. Phys. Chem. B* **2005**, *109*, 20773−20778.
(72) Baibarac, M.; mez-Romero, P. Nanocomposites Based on Conducting Polymers and Carbon Nanotubes: From Fancy Materials to Functional Applications. *J. Nanosci. Nanotechnol.* **2006**, *6*, 289−302.
(73) Pyun, J.; Matyjaszewski, K. Synthesis of Nanocomposite Organic/Inorganic Hybrid Materials Using Controlled/"Living" Radical Polymerization. *Chem. Mater.* **2001**, *13*, 3436−3448.
(74) Xiao, Y.; Li, C. M. Nanocomposites: From Fabrications to Electrochemical Bioapplications. *Electroanalysis* **2008**, *20*, 648−662.
(75) Gill, I. Bio-doped Nanocomposite Polymers: Sol−Gel Bioencapsulates. *Chem. Mater.* **2001**, *13*, 3404−3421.
(76) Wen, J.; Wilkes, G. L. Organic/Inorganic Hybrid Network Materials by the Sol−Gel Approach. *Chem. Mater.* **1996**, *8*, 1667−1681.
(77) Schmidt, H.; Jonschker, G.; Goedicke, S.; Mennig, M. The Sol-Gel Process as a Basic Technology for Nanoparticle-Dispersed Inorganic-Organic Composites. *J. Sol-Gel Sci. Technol.* **2000**, *19*, 39−51.
(78) Shchipunov, Y. A.; Karpenko, T. y. Y. Hybrid Polysaccharide−Silica Nanocomposites Prepared by the Sol−Gel Technique. *Langmuir* **2004**, *20*, 3882−3887.
(79) Nanu, M.; Schoonman, J.; Goossens, A. Inorganic Nanocomposites of n- and p-Type Semiconductors: A New Type of Three-Dimensional Solar Cell. *Adv. Mater.* **2004**, *16*, 453−456.
(80) Vu, Q.-T.; Pavlik, M.; Hebestreit, N.; Rammelt, U.; Plieth, W.; Pfleger, J. Nanocomposites based on Titanium Dioxide and Polythiophene: Structure and Properties. *React. Funct. Polym.* **2005**, *65*, 69−77.
(81) Misra, S. C. K.; Mathur, P.; Srivastava, B. K. Vacuum-Deposited Nanocrystalline Polyaniline Thin Film Sensors for Detection of Carbon Monoxide. *Sens. Actuators, A* **2004**, *114*, 30−35.
(82) Misra, S. C. K.; Mathur, P.; Yadav, M.; Tiwari, M. K.; Garg, S. C.; Tripathi, P. Preparation and Characterization of Vacuum Deposited Semiconducting Nanocrystalline Polymeric Thin Film Sensors for Detection of HCl. *Polymer* **2004**, *45*, 8623−8628.
(83) Meristoudi, A.; Athanasekos, L.; Vasileiadis, M.; Pispas, S.; Mousdis, G.; Karoutsos, E.; Alexandropoulos, D.; Du, H.; Tsigara, A.; Kibasi, K.; Perrone, A.; Vainos, N. A. Nanocomposite Hybrid Photonic Media for Remote Point Sensors. *J. Opt. A: Pure Appl. Opt.* **2009**, *11*, 034005.
(84) Kreno, L. E.; Leong, K.; Farha, O. K.; Allendorf, M.; Van Duyne, R. P.; Hupp, J. T. Metal−Organic Framework Materials as Chemical Sensors. *Chem. Rev.* **2011**, *112*, 1105−1125.
(85) Heinze, J. r.; Frontana-Uribe, B. A.; Ludwigs, S. Electrochemistry of Conducting Polymers—Persistent Models and New Concepts. *Chem. Rev.* **2010**, *110*, 4724−4771.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

(86) Rajesh Ahuja, T.; Kumar, D. Recent Progress in the Development of Nano-Structured Conducting Polymers/Nanocomposites for Sensor Applications. *Sens. Actuators, B* **2009**, *136*, 275−286.
(87) Dhand, C.; Arya, S. K.; Singh, S. P.; Singh, B. P.; Datta, M.; Malhotra, B. D. Preparation of Polyaniline/Multiwalled Carbon Nanotube Composite by Novel Electrophoretic Route. *Carbon* **2008**, *46*, 1727−1735.
(88) Shinar, R. S. J. Organic Electronics in Sensors and Biotechnology. *Electrophoretically Deposited Polymers for Organic Electronics*. McGraw-Hill Professional: New York, 2009.
(89) Zhitomirsky, I. Electrophoretic Deposition of Organic−Inorganic Nanocomposites. *J. Mater. Sci.* **2006**, *41*, 8186−8195.
(90) McDonagh, C.; Burke, C. S.; MacCraith, B. D. Optical Chemical Sensors. *Chem. Rev.* **2008**, *108*, 400−422.
(91) Homola, J.; Yee, S. S.; Gauglitz, G. Surface Plasmon Resonance Sensors: Review. *Sens. Actuators, B* **1999**, *54*, 3−15.
(92) Korotcenkov, G. *Chemical Sensors: Comprehensive Sensor Technologies*; Momentum Press: New York, 2012; Vol. 5.
(93) Stevens, N.; Akins, D. L. Dye-Doped Inorganic/Organic Composite Films as Fluorescence Sensors for Methanol Vapor. *Sens. Actuators, B* **2007**, *123*, 59−64.
(94) Homola, J. Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species. *Chem. Rev.* **2008**, *108*, 462−493.
(95) Bouvet, M. Phthalocyanine-Based Field-Effect Transistors as Gas Sensors. *Anal. Bioanal. Chem.* **2006**, *384*, 366−373.
(96) Torsi, L.; Farinola, G. M.; Marinelli, F.; Tanese, M. C.; Omar, O. H.; Valli, L.; Babudri, F.; Palmisano, F.; Zambonin, P. G.; Naso, F. A Sensitivity-Enhanced Field Effect Chiral Sensor. *Nat. Mater.* **2008**, *7*, 412−417.
(97) Arshak, K.; Moore, E.; Lyons, G.; Harris, J.; Clifford, S. A Review of Gas Sensors Employed in Electronic Nose Applications. *Sens. Rev.* **2004**, *24*, 181−198.
(98) Caliendo, C.; Verardi, P.; Verona, E.; D'Amico, A.; Natale, C. D.; Saggio, G.; Serafini, M.; Paolesse, R.; Huq, S. E. Advances in SAW-based Gas Sensors. *Smart Mater. Struct.* **1997**, *6*, 689.
(99) Lim, C.; Wang, W.; Yang, S.; Lee, K. Development of SAW-based Multi-Gas Sensor for Simultaneous Detection of CO_2 and NO_2. *Sens. Actuators, B* **2011**, *154*, 9−16.
(100) Özmen, A.; Tekce, F.; Ebeoğlu, M. A.; Taşaltın, C.; Öztürk, Z. Z. Finding the Composition of Gas Mixtures by a Phthalocyanine-Coated QCM Sensor Array and an Artificial Neural Network. *Sens. Actuators, B* **2006**, *115*, 450−454.
(101) Bakker, E. Electrochemical Sensors. *Anal. Chem.* **2004**, *76*, 3285−3298.
(102) Korotcenkov, G.; Han, S. D.; Stetter, J. R. Review of Electrochemical Hydrogen Sensors. *Chem. Rev.* **2009**, *109*, 1402−1433.
(103) Stradiotto, N. R.; Yamanaka, H.; Zanoni, M. V. B. Electrochemical Sensors: A Powerful Tool in Analytical Chemistry. *J. Br. Chem. Soc.* **2003**, *14*, 159−173.
(104) Tierney, M. J.; Kim, H. O. L. Electrochemical Gas Sensor with Extremely Fast Response Times. *Anal. Chem.* **1993**, *65*, 3435−3440.
(105) Aswal, D. K.; Gupta, S. K. *Science and technology of Chemiresistor Gas Sensors*. Nova Science Publishers: 2007.
(106) Franke, M. E.; Koplin, T. J.; Simon, U. Metal and Metal Oxide Nanoparticles in Chemiresistors: Does the Nanoscale Matter? *Small* **2006**, *2*, 36−50.
(107) Bobacka, J.; Ivaska, A.; Lewenstam, A. Potentiometric Ion Sensors. *Chem. Rev.* **2008**, *108*, 329−351.
(108) Xie, X.; Bakker, E. Non-Severinghaus Potentiometric Dissolved CO_2 Sensor with Improved Characteristics. *Anal. Chem.* **2013**, *85*, 1332−1336.
(109) Cha, W.; Tung, Y.-C.; Meyerhoff, M. E.; Takayama, S. Patterned Electrode-Based Amperometric Gas Sensor for Direct Nitric Oxide Detection within Microfluidic Devices. *Anal. Chem.* **2010**, *82*, 3300−3305.
(110) Ram, M. K.; Yavuz, O.; Aldissi, M. NO_2 Gas Sensing Based on Ordered Ultrathin Films of Conducting Polymer and its Nanocomposite. *Synth. Met.* **2005**, *151*, 77−84.
(111) Destaillats, H.; Maddalena, R. L.; Singer, B. C.; Hodgson, A. T.; McKone, T. E. Indoor Pollutants Emitted by Office Equipment: A Review of Reported Data and Information Needs. *Atmos. Environ.* **2008**, *42*, 1371−1388.
(112) Yu, C.; Crump, D. A Review of the Emission of VOCs from Polymeric Materials used in Buildings. *Build. Environ.* **1998**, *33*, 357−374.
(113) Molhave, L. Volatile Organic Compounds, Indoor Air Quality and Health. *Indoor Air* **1991**, *1*, 357−376.
(114) Hakim, M.; Broza, Y. Y.; Barash, O.; Peled, N.; Phillips, M.; Amann, A.; Haick, H. Volatile Organic Compounds of Lung Cancer and Possible Biochemical Pathways. *Chem. Rev.* **2012**, *112*, 5949−5966.
(115) Ashley, D. L.; Bonin, M. A.; Cardinali, F. L.; McCraw, J. M.; Wooten, J. V. Measurement of Volatile Organic Compounds in Human Blood. *Environ. Health Perspect.* **1996**, *104*, 871.
(116) Burn, J.; Henk, J.; Bloemen, T. *Chemistry and Analysis of Volatile Organic Compounds in the Environment*; Springer: Berlin, 1993.
(117) Ott, W. R.; Roberts, J. W. Everyday Exposure to Toxic Pollutants. *Sci. Am.* **1998**, *278*, 86−91.
(118) Tchobanoglous, G.; Burton, F. L. Wastewater Engineering. *Management* **1991**, *7*, 1−4.
(119) Potyrailo, R. A.; Leach, A. M.; Surman, C. M. Multisize CdSe Nanocrystal/Polymer Nanocomposites for Selective Vapor Detection Identified from High-Throughput Screening Experimentation. *ACS Combi. Sci.* **2012**, *14*, 170−178.
(120) Li, W.; Kim, D. Polyaniline/Multiwall Carbon Nanotube Nanocomposite for Detecting Aromatic Hydrocarbon Vapors. *J. Mater. Sci.* **2011**, *46*, 1857−1861.
(121) Campos, M.; Bulhões, L. O. S.; Lindino, C. A. Gas-sensitive Characteristics of Metal/Semiconductor Polymer Schottky Device. *Sens. Actuators. A* **2000**, *87*, 67−71.
(122) Sadek, A. Z.; Wlodarski, W.; Shin, K.; Kaner, R. B.; Kalantar-zadeh, K. A Polyaniline/WO_3 Nanofiber Composite Based ZnO/64° YX $LiNbO_3$ SAW Hydrogen Gas Sensor. *Synth. Met.* **2008**, *158*, 29−32.
(123) He, J. H.; Ho, C. H.; Chen, C. Y. Polymer Functionalized ZnO Nanobelts as Oxygen Sensors with a Significant Response Enhancement. *Nanotechnology* **2009**, *20*, 065503.
(124) Arshak, K.; Gaidan, I. Development of an Array of Polymer/MnO_2/Fe_2O_3 Mixtures for use in Gas Sensing Applications. *Sens. Actuators, B* **2006**, *118*, 386−392.
(125) Arshak, K.; Gaidan, I. NiO/Fe_2O_3 Polymer Thick Films as Room Temperature Gas Sensors. *Thin Solid Films* **2006**, *495*, 286−291.
(126) Vaddiraju, S.; Gleason, K. K. Selective Sensing Of Volatile Organic Compounds using Novel Conducting Polymer−Metal Nanoparticle Hybrids. *Nanotechnology* **2010**, *21*, 125503.
(127) Sayago, I.; Fernández, M. J.; Fontecha, J. L.; Horrillo, M. C.; Vera, C.; Obieta, I.; Bustero, I. New Sensitive Layers for Surface Acoustic Wave Gas Sensors Based on Polymer and Carbon Nanotube Composites. *Sens. Actuators, B* **2012**, *175*, 67−72.
(128) Tatavarty, R.; Hwang, E. T.; Park, J.-W.; Kwak, J.-H.; Lee, J.-O.; Gu, M. B. Conductive Quantum Dot-Encapsulated Electrospun Nanofibers from Polystyrene and Polystyrene-Co-Maleic Anhydride Copolymer Blend as Gas Sensors. *React.Funct. Polym.* **2011**, *71*, 104−108.
(129) Mondal, S. P.; Bera, S.; Narender, G.; Ray, S. K. CdSe Quantum Dots-Poly(3-Hexylthiophene) Nanocomposite Sensors for Selective Chloroform Vapor Detection at Room Temperature. *Appl. Phys. Lett.* **2012**, *101*, 173108.
(130) Sharma, S.; Nirkhe, C.; Pethkar, S.; Athawale, A. A. Chloroform Vapour Sensor based on Copper/Polyaniline Nanocomposite. *Sens. Actuators, B* **2002**, *85*, 131−136.
(131) Ma, X.; Wang, M.; Li, G.; Chen, H.; Bai, R. Preparation of Polyaniline−TiO_2 Composite Film with in Situ Polymerization Approach and its Gas-Sensitivity at Room Temperature. *Mater. Chem. Phys.* **2006**, *98*, 241−247.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

(132) Hierlemann, A.; Bodenhöfer, K.; Fluck, M.; Schurig, V.; Göpel, W. Selective Detection of Nitrogen and Oxygen Containing Volatile Organic Compounds: Use Of Metal-Modified Polysiloxanes As Sensor Coatings. *Anal. Chim. Acta* **1997**, *346*, 327−339.
(133) Wang, Y. D.; Djerdj, I.; Antonietti, M.; Smarsly, B. Polymer-Assisted Generation of Antimony-Doped SnO_2 Nanoparticles with High Crystallinity for Application in Gas Sensors. *Small* **2008**, *4*, 1656−1660.
(134) Wang, J.; Matsubara, I.; Murayama, N.; Woosuck, S.; Izu, N. The Preparation of Polyaniline Intercalated MoO3 Thin Film and its Sensitivity to Volatile Organic Compounds. *Thin Solid Films* **2006**, *514*, 329−333.
(135) Itoh, T.; Matsubara, I.; Shin, W.; Izu, N. Synthesis and Characterization of Layered Organic/Inorganic Hybrid Thin Films Based on Molybdenum Trioxide with Poly(N-methylaniline) for VOC sensor. *Mater. Lett.* **2007**, *61*, 4031−4034.
(136) Itoh, T.; Matsubara, I.; Shin, W.; Izu, N.; Nishibori, M. Preparation of Layered Organic−Inorganic Nanohybrid Thin Films of Molybdenum Trioxide with Polyaniline Derivatives for Aldehyde Gases Sensors of Several Tens ppb Level. *Sens. Actuators, B* **2008**, *128*, 512−520.
(137) Lu, J.; Park, B. J.; Kumar, B.; Castro, M.; Choi, H. J.; Feller, J.-F. Polyaniline Nanoparticle−Carbon Nanotube Hybrid Network Vapour Sensors with Switchable Chemo-Electrical Polarity. *Nanotechnology* **2010**, *21*, 255501.
(138) de Julián Fernández, C.; Manera, M. G.; Spadavecchia, J.; Maggioni, G.; Quaranta, A.; Mattei, G.; Bazzan, M.; Cattaruzza, E.; Bonafini, M.; Negro, E.; Vomiero, A.; Carturan, S.; Scian, C.; Mea, G. D.; Rella, R.; Vasanelli, L.; Mazzoldi, P. Study of the Gas Optical Sensing Properties of Au-Polyimide Nanocomposite Films Prepared by Ion Implantation. *Sens. Actuators, B* **2005**, *111*−*112*, 225−229.
(139) Savage, N. O. Gas Sensing Composites of Metal Oxides with Vapor-Deposited Polypyrrole. *Sens. Actuators, B* **2009**, *143*, 6−11.
(140) Barkade, S. S.; Naik, J. B.; Sonawane, S. H. Ultrasound Assisted Miniemulsion Synthesis of Polyaniline/Ag Nanocomposite and its Application for Ethanol Vapor Sensing. *Colloids Surf., A* **2011**, *378*, 94−98.
(141) Kumar, B.; Feller, J.-F.; Castro, M.; Lu, J. Conductive Bio-Polymer Nano-Composites (CPC): Chitosan-Carbon Nanotube Transducers Assembled via Spray Layer-by-Layer for Volatile Organic Compound Sensing. *Talanta* **2010**, *81*, 908−915.
(142) Athawale, A. A.; Bhagwat, S. V.; Katre, P. P. Nanocomposite of Pd−Polyaniline as a Selective Methanol Sensor. *Sens. Actuators, B* **2006**, *114*, 263−267.
(143) Han, L.; Andrady, A. L.; Ensor, D. S. Chemical Sensing Using Electrospun Polymer/Carbon Nanotube Composite Nanofibers with Printed-on Electrodes. *Sens. Actuators, B* **2013**, *186*, 52−55.
(144) Feller, J. F.; Lu, J.; Zhang, K.; Kumar, B.; Castro, M.; Gatt, N.; Choi, H. J. Novel Architecture of Carbon Nanotube Decorated Poly(Methyl Methacrylate) Microbead Vapour Sensors Assembled by Spray Layer by Layer. *J. Mater. Chem.* **2011**, *21*, 4142−4149.
(145) Lvova, L.; Mastroianni, M.; Pomarico, G.; Santonico, M.; Pennazza, G.; Di Natale, C.; Paolesse, R.; D'Amico, A. Carbon Nanotubes Modified with Porphyrin Units for Gaseous Phase Chemical Sensing. *Sens. Actuators, B* **2012**, *170*, 163−171.
(146) Evans, C. D.; Monteith, D. T.; Fowler, D.; Cape, J. N.; Brayshaw, S. Hydrochloric Acid: An Overlooked Driver of Environmental Change. *Environ. Sci. Technol.* **2011**, *45*, 1887−1894.
(147) Tripathy, S. K.; Woo, J. Y.; Han, C.-S. Highly Selective Colorimetric Detection of Hydrochloric Acid Using Unlabeled Gold Nanoparticles and an Oxidizing Agent. *Anal. Chem.* **2011**, *83*, 9206−9212.
(148) Jang, J.; Bae, J. Carbon Nanofiber/Polypyrrole Nanocable as Toxic Gas Sensor. *Sens. Actuators, B* **2007**, *122*, 7−13.
(149) Cano, M.; Castillero, P.; Roales, J.; Pedrosa, J. M.; Brittle, S.; Richardson, T.; González-Elipe, A. R.; Barranco, A. A transparent TMPyP/TiO_2 Composite Thin Film as an HCl Sensitive Opto-chemical Gas Sensor. *Sens. Actuators, B* **2010**, *150*, 764−769.
(150) Timmer, B. r.; Olthuis, W.; Berg, A. v. d. Ammonia Sensors and Their Applications-A Review. *Sens. Actuators, B* **2005**, *107*, 666−677.
(151) Kowalchuk, G. A.; Stephen, J. R. Ammonia-oxidizing Bacteria: A Model for Molecular Microbial Ecology. *Annu. Rev. Microbiol.* **2001**, *55*, 485−529.
(152) Kristensen, H.; Wathes, C. Ammonia and Poultry Welfare: A review. *World's Poultry Sci. J.* **2000**, *56*, 235−245.
(153) Venditti, I.; Fratoddi, I.; Russo, M. V.; Bearzotti, A. A Nanostructured Composite Based on Polyaniline and Gold Nanoparticles: Synthesis and Gas Sensing Properties. *Nanotechnology* **2013**, *24*, 155503.
(154) Lobotka, P.; Kunzo, P.; Kovacova, E.; Vavra, I.; Krizanova, Z.; Smatko, V.; Stejskal, J.; Konyushenko, E. N.; Omastova, M.; Spitalsky, Z.; Micusik, M.; Krupa, I. Thin Polyaniline and Polyaniline/Carbon Nanocomposite Films for Gas Sensing. *Thin Solid Films* **2011**, *519*, 4123−4127.
(155) Tai, H.; Jiang, Y.; Xie, G.; Yu, J.; Chen, X. Fabrication and Gas Sensitivity of Polyaniline−Titanium Dioxide Nanocomposite Thin Film. *Sens. Actuators, B* **2007**, *125*, 644−650.
(156) Tai, H.; Jiang, Y.; Xie, G.; Yu, J. Preparation, Characterization and Comparative NH3-sensing Characteristic Studies of PANI/Inorganic Oxides Nanocomposite Thin Films. *J. Mater. Sci. Technol.* **2010**, *26*, 605−613.
(157) Van Hieu, N.; Dung, N. Q.; Tam, P. D.; Trung, T.; Chien, N. D. Thin Film Polypyrrole/SWCNTs Nanocomposites Based NH_3 Sensor Operated at Room Temperature. *Sens. Actuators, B* **2009**, *140*, 500−507.
(158) Hong, L.; Li, Y.; Yang, M. Fabrication and Ammonia Gas Sensing of Palladium/Polypyrrole Nanocomposite. *Sens. Actuators, B* **2010**, *145*, 25−31.
(159) Lim, J.-H.; Phiboolsirichit, N.; Mubeen, S.; Deshusses, M. A.; Mulchandani, A.; Myung, N. V. Electrical and Gas Sensing Properties of Polyaniline Functionalized Single-Walled Carbon Nanotubes. *Nanotechnology* **2010**, *21*, 075502.
(160) Jian, J.; Guo, X.; Lin, L.; Cai, Q.; Cheng, J.; Li, J. Gas-Sensing Characteristics of Dielectrophoretically Assembled Composite Film of Oxygen Plasma-Treated SWCNTs and PEDOT/PSS Polymer. *Sens. Actuators, B* **2013**, *178*, 279−288.
(161) Jiang, S.; Chen, J.; Tang, J.; Jin, E.; Kong, L.; Zhang, W.; Wang, C. Au Nanoparticles-Functionalized Two-Dimensional Patterned Conducting PANI Nanobowl Monolayer for Gas Sensor. *Sens. Actuators, B* **2009**, *140*, 520−524.
(162) Chang, Q.; Zhao, K.; Chen, X.; Li, M.; Liu, J. Preparation of Gold/Polyaniline/Multiwall Carbon Nanotube Nanocomposites and Application in Ammonia Gas Detection. *J. Mater. Sci.* **2008**, *43*, 5861−5866.
(163) Zhang, T.; Mubeen, S.; Yoo, B.; Myung, N. V.; Deshusses, M. A. A gas Nanosensor Unaffected by Humidity. *Nanotechnology* **2009**, *20*, 255501.
(164) Gong, J.; Li, Y.; Hu, Z.; Zhou, Z.; Deng, Y. Ultrasensitive NH_3 Gas Sensor from Polyaniline Nanograin Enchased TiO_2 Fibers. *J. Phys. Chem. C* **2010**, *114*, 9970−9974.
(165) Tao, W.-H.; Tsai, C.-H. H_2S Sensing Properties of Noble Metal Doped WO_3 Thin Film Sensor Fabricated by Micromachining. *Sens. Actuators, B* **2002**, *81*, 237−247.
(166) Geng, L.; Wang, S.; Zhao, Y.; Li, P.; Zhang, S.; Huang, W.; Wu, S. Study of the Primary Sensitivity of Polypyrrole/r-Fe_2O_3 to Toxic Gases. *Mater. Chem. Phys.* **2006**, *99*, 15−19.
(167) Raut, B. T.; Chougule, M. A.; Nalage, S. R.; Dalavi, D. S.; Mali, S.; Patil, P. S.; Patil, V. B. CSA doped Polyaniline/CdS Organic−Inorganic Nanohybrid: Physical and Gas Sensing Properties. *Ceram. Int.* **2012**, *38*, 5501−5506.
(168) Shirsat, M. D.; Bangar, M. A.; Deshusses, M. A.; Myung, N. V.; Mulchandani, A. Polyaniline Nanowires-Gold Nanoparticles Hybrid Network based Chemiresistive Hydrogen Sulfide Sensor. *Appl. Phys. Lett.* **2009**, *94*, 083502−3.
(169) Starke, T. K. H.; Coles, G. S. V.; Ferkel, H. High Sensitivity NO_2 Sensors for Environmental Monitoring Produced Using Laser

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Ablated Nanocrystalline Metal Oxides. *Sens. Actuators, B* **2002**, *85*, 239−245.
(170) Fleming, W. J. Overview of Automotive Sensors. *Sens. J., IEEE* **2001**, *1*, 296−308.
(171) Menil, F.; Coillard, V.; Lucat, C. Critical Review of Nitrogen Monoxide Sensors for Exhaust Gases of Lean Burn Engines. *Sens. Actuators, B* **2000**, *67*, 1−23.
(172) Prakash, S.; Rajesh, S.; Singh, S. R.; Karunakaran, C.; Vasu, V. Electrochemical Incorporation of Hemin in a ZnO−PPy Nanocomposite on a Pt Electrode as NOx Sensor. *Analyst* **2012**, *137*, 5874.
(173) Geng, L.; Zhao, Y.; Huang, X.; Wang, S.; Zhang, S.; Huang, W.; Wu, S. The Preparation and Gas Sensitivity Study of Polypyrrole/Zinc Oxide. *Synth. Met.* **2006**, *156*, 1078−1082.
(174) Xu, M.; Zhang, J.; Wang, S.; Guo, X.; Xia, H.; Wang, Y.; Zhang, S.; Huang, W.; Wu, S. Gas Sensing Properties of SnO_2 Hollow Spheres/Polythiophene Inorganic−Organic Hybrids. *Sens. Actuators, B* **2010**, *146*, 8−13.
(175) Sadek, A. Z.; Wlodarski, W.; Shin, K.; Kaner, R. B.; Kalantar-zadeh, K. A Layered Surface Acoustic Wave Gas Sensor Based on a Polyaniline/In_2O_3 Nanofibre Composite. *Nanotechnology* **2006**, *17*, 4488−4492.
(176) Do, J.-S.; Chang, W.-B. Amperometric Nitrogen Dioxide Gas Sensor Based on Pan/Au/Nafion Prepared by Constant Current and Cyclic Voltammetry Methods. *Sens. Actuators, B* **2004**, *101*, 97−106.
(177) Hoherčáková, Z.; Opekar, F. Au/PVC Composite-A new Material for Solid-State Gas Sensors: Detection of Nitrogen Dioxide in the Air. *Sens. Actuators, B* **2004**, *97*, 379−386.
(178) Nguyen, D. T.; Nguyen, M. T.; Ho, G. T.; Nguyen, T. N.; Reisberg, S.; Piro, B.; Pham, M. C. Design of Interpenetrated Network MWCNT/Poly(1,5-DAN) on Interdigital Electrode: Toward NO_2 Gas Sensing. *Talanta* **2013**, *115*, 713−717.
(179) Gedney, N.; Cox, P.; Betts, R.; Boucher, O.; Huntingford, C.; Stott, P. Detection of a Direct Carbon Dioxide Effect in Continental River Runoff Records. *Nature* **2006**, *439*, 835−838.
(180) Neethirajan, S.; Jayas, D.; Sadistap, S. Carbon Dioxide (CO_2) Sensors for the Agri-Food Industry-A Review. *Food Bioprocess Technol.* **2009**, *2*, 115−121.
(181) Star, A.; Han, T. R.; Joshi, V.; Gabriel, J. C.; Gruner, G. Nanoelectronic Carbon Dioxide Sensors. *Adv. Mater.* **2004**, *16*, 2049−2052.
(182) Kim, J. H.; Lee, Y. M. Gas Permeation Properties of Poly(Amide-6-B-Ethylene Oxide)−Silica Hybrid Membranes. *J. Membr. Sci.* **2001**, *193*, 209−225.
(183) Rodenas, T.; Luz, I.; Prieto, G.; Seoane, B.; Miro, H.; Corma, A.; Kapteijn, F.; L. i. Xamena, F. X.; Gascon, J. Metal-Organic Framework Nanosheets in Polymer Composite Materials for Gas Separation. *Nat. Mater.* **2015**, *14*, 48−55.
(184) Kang, W. R.; Lee, A. S.; Park, S.; Park, S.-H.; Baek, K.-Y.; Lee, K. B.; Lee, S.-H.; Lee, J.-H.; Hwang, S. S.; Lee, J. S. Free-Standing, Polysilsesquioxane-Based Inorganic/Organic Hybrid Membranes for Gas Separations. *J. Membr. Sci.* **2015**, *475*, 384−394.
(185) Raub, J. A.; Mathieu-Nolf, M.; Hampson, N. B.; Thom, S. R. Carbon Monoxide Poisoning-A Public Health Perspective. *Toxicology* **2000**, *145*, 1−14.
(186) Ernst, A.; Zibrak, J. D. Carbon Monoxide Poisoning. *N. Eng. J. Med.* **1998**, *339*, 1603−1608.
(187) Densakulprasert, N.; Wannatong, L.; Chotpattananont, D.; Hiamtup, P.; Sirivat, A.; Schwank, J. Electrical Conductivity of Polyaniline/Zeolite Composites and Synergetic Interaction with CO. *Mater. Sci. Eng., B* **2005**, *117*, 276−282.
(188) Ram, M. K.; Yavuz, O.; Lahsangah, V.; Aldissi, M. CO Gas Sensing from Ultrathin Nano-Composite Conducting Polymer Film. *Sens. Actuators, B* **2005**, *106*, 750−757.
(189) Jin, G.-P.; Peng, X.; Ding, Y.-F.; Liu, W.-Q.; Ye, J.-M. Electrodeposition of Platinum−Nickel Alloy Nanocomposites on Polyaniline-Multiwalled Carbon Nanotubes for Carbon Monoxide Redox. *J. Solid State Electrochem.* **2009**, *13*, 967−973.
(190) Chen, Z.; Lu, C. Humidity Sensors: A Review of Materials and Mechanisms. *Sens. Lett.* **2005**, *3*, 274−295.
(191) Lee, C.-Y.; Lee, G.-B. Humidity Sensors: A Review. *Sens. Lett.* **2005**, *3*, 1−15.
(192) Rittersma, Z. M. Recent achievements in Miniaturised Humidity Sensors-A Review of Transduction Techniques. *Sens. Actuators, A* **2002**, *96*, 196−210.
(193) Kumar, A.; Zhang, P.; Vincent, A.; McCormack, R.; Kalyanaraman, R.; Cho, H. J.; Seal, S. Hydrogen Selective Gas Sensor in Humid Environment Based on Polymer Coated Nanostructured Doped Tin Oxide. *Sens. Actuators, B* **2011**, *155*, 884−892.
(194) Mickelson, W.; Sussman, A.; Zettl, A. Low-Power, Fast, Selective Nanoparticle Based Hydrogen Sulfide Gas Sensor. *Appl. Phys. Lett.* **2012**, *100*, 173110.
(195) Wu, T.-T.; Chen, Y.-Y.; Chou, T.-H. A High Sensitivity Nanomaterial Based SAW Humidity Sensor. *J. Phys. D: Appl. Phys.* **2008**, *41*, 085101.
(196) Tandon, R. P.; Tripathy, M. R.; Arora, A. K.; Hotchandani, S. Gas and Humidity Response of Iron Oxide-Polypyrrole Nanocomposites. *Sens. Actuators, B* **2006**, *114*, 768−773.
(197) Power, A. C.; Betts, A. J.; Cassidy, J. F. Silver Nanoparticle Polymer Composite Based Humidity Sensor. *Analyst* **2010**, *135*, 1645−1652.
(198) Su, P.-G.; Huang, L.-N. Humidity Sensors Based on TiO_2 Nanoparticles/Polypyrrole Composite Thin Films. *Sens. Actuators, B* **2007**, *123*, 501−507.
(199) Su, P.-G.; Wang, C.-P. Flexible Humidity Sensor based on TiO_2 Nanoparticles-Polypyrrole-Poly-[3-(Methacrylamino)Propyl] Trimethyl Ammonium Chloride Composite Materials. *Sens. Actuators, B* **2008**, *129*, 538−543.
(200) Patil, D.; Seo, Y.-K.; Hwang, Y. K.; Chang, J.-S.; Patil, P. Humidity Sensing Properties of Poly(o-anisidine)/WO_3 Composites. *Sens. Actuators, B* **2008**, *128*, 374−382.
(201) Parvatikar, N.; Jain, S.; Khasim, S.; Revansiddappa, M.; Bhoraskar, S. V.; Prasad, M. V. N. A. Electrical and Humidity Sensing Properties of Polyaniline/WO_3 Composites. *Sens. Actuators, B* **2006**, *114*, 599−603.
(202) Su, P.-G.; Huang, S.-C. Humidity Sensing and Electrical Properties of a Composite Material of SiO_2 and poly-[3-(methacrylamino)propyl] Trimethyl Ammonium Chloride. *Sens. Actuators, B* **2005**, *105*, 170−175.
(203) Litao, L.; Xiongying, Y.; Kang, W.; Zhaoying, Z.; Dongjin, L.; Tianhong, C. Humidity Sensitivity of Carbon Nanotube and Poly (Dimethyldiallylammonium Chloride) Composite Films. *Sens. J., IEEE* **2009**, *9*, 1308−1314.
(204) Li, Y.; Yang, M. J.; She, Y. Humidity Sensors Using In Situ Synthesized Sodium Polystyrenesulfonate/ZnO Nanocomposites. *Talanta* **2004**, *62*, 707−712.
(205) Fuke, M. V.; Kanitkar, P.; Kulkarni, M.; Kale, B. B.; Aiyer, R. C. Effect of Particle Size Variation of Ag Nanoparticles in Polyaniline Composite on Humidity Sensing. *Talanta* **2010**, *81*, 320−326.
(206) Liu, L.; Ye, X.; Wu, K.; Zhou, Z.; Lee, D.; Cui, T. Humidity Sensitivity of Carbon Nanotube and Poly (Dimethyldiallylammonium Chloride) Composite Films. *Sens. J., IEEE* **2009**, *9*, 1308−1314.
(207) Timm, R. A.; Falla, M. P. H.; Huila, M. F. G.; Peres, H. E. M.; Ramirez-Fernandez, F. J.; Araki, K.; Toma, H. E. Vanadium oxide-Porphyrin Nanocomposites as Gas Sensor Interfaces for Probing Low Water Content in Ethanol. *Sens. Actuators, B* **2010**, *146*, 61−68.
(208) Arshak, K.; Adley, C.; Moore, E.; Cunniffe, C.; Campion, M.; Harris, J. Characterisation of Polymer Nanocomposite Sensors for Quantification of Bacterial Cultures. *Sens. Actuators, B* **2007**, *126*, 226−231.
(209) Zhan, S.; Li, D.; Liang, S.; Chen, X.; Li, X. A Novel Flexible Room Temperature Ethanol Gas Sensor based on SnO_2 Doped Poly-Diallyldimethylammonium Chloride. *Sensors* **2015**, *13*, 4378−4389.
(210) Kruefu, V.; Wisitsoraat, A.; Tuantranont, A.; Phanichphant, S. C_2H_5OH Gas Sensing Based on Poly(3-hexylthiophene)/Nb-Loaded ZnO Nanocomposite Films. *Mol. Cryst. Liq. Cryst.* **2014**, *599*, 1−7.
(211) Nasirian, S.; Moghaddam, H. M. Hydrogen Gas Sensing Based on Polyaniline/Anatase Titania Nanocomposite. *Int. J. Hydrogen Energy* **2014**, *39*, 630−642.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

(212) Tai, H.; Li, X.; Jiang, Y.; Xie, G.; Du, X. The Enhanced Formaldehyde-Sensing Properties of P3HT-ZnO Hybrid Thin Film OTFT Sensor and Further Insight into its Stability. *Sensors* **2015**, *15*, 2086−2103.
(213) Murugan, C.; Subramanian, E.; Padiyan, D. P. Enhanced Sensor Functionality of In-Situ Synthesized Polyaniline-SnO_2 Hybrids Toward Benzene and Toluene Vapors. *Sens. Actuators, B* **2014**, *205*, 74−81.
(214) Murugan, C.; Subramanian, E.; Padiyan, D. P. p-n Heterojunction Formation in Polyaniline-SnO_2 Organic-Inorganic Hybrid Composite Materials Leading to Enhancement in Sensor Functionality Toward Benzene and Toluene Vapors at Room Temperature. *Synth. Met.* **2014**, *192*, 106−112.
(215) Nia, P. M.; Lorestani, F.; Meng, W. P.; Alias, Y. A Novel Non-enzymatic H_2O_2 Sensor Based on Polypyrrole Nanofibers-Silver Nanoparticles Decorated Reduced Graphene Oxide Nanocomposites. *Appl. Surf. Sci.* **2015**, *332*, 648−656.
(216) Bochenkov, V. E.; Stephan, N.; Brehmer, L.; Zagorskii, V. V.; Sergeev, G. B. Sensor Activity of Thin Polymer Films Containing Lead Nanoparticles. *Colloid. Surf., A* **2002**, *198−200*, 911−915.
(217) Talwar, V.; Singh, O.; Singh, R. C. ZnO Assisted Polyaniline Nanofibers and its Application as Ammonia Gas Sensor. *Sens. Actuators, B* **2014**, *191*, 276−282.
(218) Wang, L.; Huang, H.; Xiao, S.; Cai, D.; Liu, Y.; Liu, B.; Wang, D.; Wang, C.; Li, H.; Wang, Y.; Li, Q.; Wang, T. Enhanced Sensitivity and Stability of Room-Temperature NH3 Sensors Using Core-Shell CeO2 Nanoparticles@Cross-linked PANI with p-n Heterojunctions. *ACS Appl. Mater. Interface* **2014**, *6*, 14131−14140.
(219) Pandey, S.; Goswami, G. K.; Nanda, K. K. Nanocomposite Based Flexible Ultrasensitive Resistive Gas Sensor for Chemical Reactions Studies. *Sci. Rep.* **2015**, *3*.
(220) Su, P.-G.; Peng, Y.-T. Fabrication of a Room-Temperature H_2S Gas Sensor Based on PPy/WO_3 Nanocomposite Films by In-Situ Photopolymerization. *Sens. Actuators, B* **2014**, *193*, 637−643.
(221) Bai, S.; Zhang, K.; Sun, J.; Zhang, D.; Luo, R.; Li, D.; Liu, C. Polythiophene-WO_3 Hybrid Architectures For Low-Temperature H_2S Detection. *Sens. Actuators, B* **2014**, *197*, 142−148.
(222) Fedoruk, G. G.; Sagaidak, D. I.; Misevich, A. V.; Pochtenny, A. E. Electrical and Gas Sensing Properties of Copper Phthalocyanine−Polymer Composites. *Sens. Actuators, B* **1998**, *48*, 351−355.
(223) Tiwari, A.; Gong, S. Electrochemical Synthesis of Chitosan-co-polyaniline/WO_3·nH_2O Composite Electrode for Amperometric Detection of NO_2 Gas. *Electroanalysis* **2008**, *20*, 1775−1781.
(224) Nalage, S. R.; Mane, A. T.; Pawar, R. C.; Lee, C. S.; Patil, V. B. Polypyrrole-NiO Hybrid Nanocomposite Films: Highly Selective, Sensitive, and Reproducible NO_2 Sensors. *Ionics* **2014**, *20*, 1607−1616.
(225) Navale, S.; Khuspe, G.; Chougule, M.; Patil, V. Room Temperature NO_2 Gas Sensor Based on PPy/α- Fe_2O_3 Hybrid Nanocomposites. *Ceram. Int.* **2014**, *40*, 8013−8020.
(226) Mane, A.; Navale, S.; Sen, S.; Aswal, D.; Gupta, S.; Patil, V. Nitrogen Dioxide (NO_2) Sensing Performance of P-Polypyrrole/N-Tungsten Oxide Hybrid Nanocomposites at Room temperature. *Org. Electron.* **2015**, *16*, 195−204.
(227) Tsigara, A.; Mountrichas, G.; Gatsouli, K.; Nichelatti, A.; Pispas, S.; Madamopoulos, N.; Vainos, N. A.; Du, H. L.; Roubani-Kalantzopoulou, F. Hybrid Polymer/Cobalt Chloride Humidity Sensors Based on Optical Diffraction. *Sens. Actuators, B* **2007**, *120*, 481−486.

DOI: 10.1021/cr400659h
Chem. Rev. XXXX, XXX, XXX−XXX

Microsyst Technol
DOI 10.1007/s00542-015-2557-9

TECHNICAL PAPER

Oxide bound impact on hot-carrier degradation for gate electrode workfunction engineered (GEWE) silicon nanowire MOSFET

Neha Gupta[1] · Ajay Kumar[1] · Rishu Chaujar[1]

Received: 29 October 2014 / Accepted: 20 April 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract In this present work, we explore the hot carrier fidelity of gate electrode workfunction engineered silicon nanowire (GEWE-SiNW) MOSFET at 300 K using DEVEDIT-3D device editor and ATLAS device simulation software. TCAD simulation shows reduction in the hot carrier reliability of a GEWE SiNW MOSFET in terms of electron temperature, electron velocity and Hot Electron gate current for reflecting its efficacy in high power CMOS applications. Further, a comparative investigation for different values of oxide thickness and high-k has been done to analyze the performance of GEWE-SiNW MOSFET in terms of electrical parameters such as conduction band, DIBL, electric field, electron temperature, electric velocity and gate current. It has been clearly shown that with oxide thickness 0.5 nm the hot-carrier reliability and device performance improves in comparison to oxide thickness 2.5 nm. In addition, with k = 21(HfO_2) device performance in terms of hot-carrier reliability further enhanced due to increased capacitance and thus offer its effectiveness in sub-nm range analog applications.

✉ Rishu Chaujar
rishu.phy@dce.edu

Neha Gupta
nehagupta_dtu@yahoo.com

Ajay Kumar
ajaykumar@dtu.ac.in

[1] Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

1 Introduction

From past few decades, the major concern for conventional CMOS devices were scaling effects such as leakage current, hot carrier effects, DIBL and current driving ability. To overcome such effects a number of device engineering schemes such as source/drain engineering (Kranti et al. 2006), metal gate workfunction engineering (Long et al. 1997; Deb et al. 2012; Dubey et al. 2013), channel engineering (Yu et al. 1997) and some novel device structures such as Double gate MOSFET (Barsan 1981; Venkatesan et al. 1992), cylindrical gate MOSFET (Sharma and Vishvakarma 2013), Fin-FETs (Sohn and Kang 2012), Nanowire MOSFET (Iwai et al. 2011; Wang et al. 2007) have been explored and employed during the past few years. Among them, due to ideal symmetry, excellent gate controllability and potential benefits of silicon nanowire, gate-all-around (GAA) silicon nanowire MOSFET (Chen and Tan 2014; Wang et al. 2013) is emerging as a most promising candidate for future silicon based nano-devices. This structure manifest a superior control of short channel effects because of excellent electrostatic coupling between the conducting channel and the surrounding gate. The nanowire (NW) transistors can be seen as the utmost integration of the innovative nano-devices and it is one of the most promising means of overwhelm the limits of planar silicon devices because of their suitability with gate all around design. Also, from device and circuit developers they have procured remarkable attention because of its potential for high performance and building highly dense electronic circuits. Current approaches in nanoscale fabrication techniques have shown that semiconductor nanowires may turn into a candidate for next generation technologies. Silicon and Germanium nanowire transistors are also significant because of their adaptability with the CMOS technology

Published online: 07 May 2015

(Mukherjee and Maiti 2012). Short-channel effects (SCEs) result in many problems such as increase in power consumption and leakage current (De and Borkar 1999). Hot-electron degradation is another unenviable effect of short-channel devices (Orouji and Kumar 2005; Krieger et al. 2001). As the scaling of transistors penetrates into the sub-micron regime, the electric field grows and causes electrons and holes to gain high kinetic energy, referred to as "hot carriers". These hot carriers have ample of high energies and thrust to allow them to be injected from the semiconductor into the surrounding oxide films such as the gate and sidewall silicon oxides (Groeseneken 2001). The existence of such free mobile carriers in the oxide provokes numerous physical devastation processes that can highly change the device characteristics over prolonged periods. The stack of damage ultimately causes the circuit to fail as decisive parameters such as threshold voltage shift due to aforesaid damage. The accumulation of damage resulting degradation in device behavior due to hot carrier injection is called hot carrier degradation (Pan et al. 1994; Kaur et al. 2007).

Drain induced barrier lowering (DIBL) is an important parameter to examine SCEs in VLSI MOSFET devices (Arora 2012). It is the variation in threshold voltage at different ($\Delta V_{th}/\Delta V_{DS}$) drain to source bias. The variation of the bending energy gap due to the threshold voltage variation is mainly governed by the increase of the drain voltage (Sharma 2012). The high electric field due to a higher V_{ds} enter deeper into the conducting channel and causes more pronounced short channel effect. DIBL tunneling affects the I–V characteristic in sub-threshold regime by a lateral shift it can be determined by the percentage of the shift in barrier height value to the drain voltage difference (Jeon et al. 2009) As the issue of hot carrier and DIBL becomes increasingly important in deep-submicron MOSFETs, it is thus increasingly essential to analyze the hot carrier reliability in terms of hot-electron-injected gate current, impact-ionization substrate current, and electron velocity and temperature near the drain end for their impact on the overall chip performance (Chaujar et al. 2008).

Moreover, the downscaling of MOSFETs dimensions to sub-nm will face the challenges when SiO_2 is used as gate dielectric. When oxide thickness of gate oxide is below 1.5 nm, the electrons tunnel through the gate oxide easily, this increases leakage current and thus leads to excessive power consumption and reduced device reliability. Therefore, it is required for high-k materials to replace SiO_2 as a gate dielectric in order to reduce the high standby power (Momose et al. 1996; Nirmal et al. 2013; Sinha and Chaudhury 2013). The work thus improves and envisions the feasibility of a GEWE-SiNW MOSFET (Gupta and Chaujar 2014) for high-performance applications where hot carrier reliability is a primary concern. TCAD simulations using ATLAS demonstrate that the GEWE-SiNW MOSFET structure exhibits significantly improved hot carrier performance, where low power consumption is required and in digital logic and memory applications where fast switching action of MOS is needed. In addition, the influence of oxide thickness variation and efficacy of different high-k materials on performance of GEWE-SiNW MOSFET has also been discussed in this paper.

The complete work in this paper is systematized as follow: Sect. 2 elucidate the 3D-device architecture and its default parameters used in this work along with simulation models. In Sect. 3, first we calibrate our simulation models with experimental results and then we compare the device performance in terms of hot-carrier reliability and further we examine the impact of thinning of gate oxide thickness and high-K of GEWE-SiNW MOSFET on various parameters including Hot-Carrier Metrics such as HE-Ig, Electron Temperature, electron velocity and DIBL for the n-MOSFET. Finally the conclusions are drawn in Sect. 4.

2 Device structure: parameters and simulation models

Figure 1 shows the simulated device structure of GEWE SiNW MOSFET. All simulations have been performed using ATLAS and DEVEDIT 3D device simulator. Drift–diffusion model is not suitable for sub-50-nm gate length devices, since it neglects the nonlocal effects. In our simulation, we have adopted the hydrodynamic energy transport model which includes the continuity equations, momentum transport equations, energy balance equations of the carriers, and the Poisson equation. It can model the nonlocal transport phenomenon and is more accurate than the drift–diffusion method.

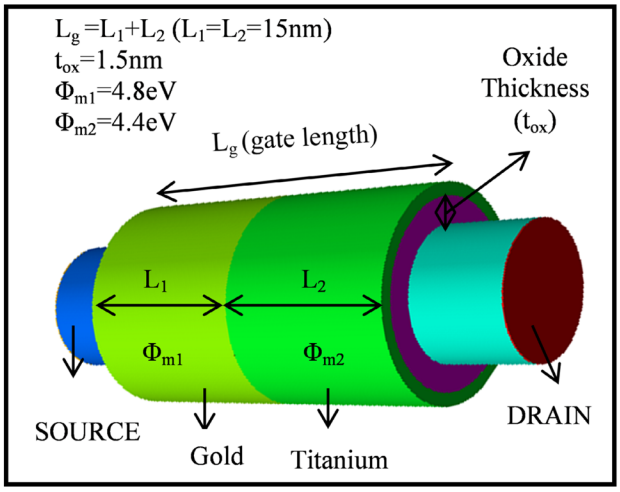


Fig. 1 Simulated device structure: GEWE SiNW MOSFET

The Energy Balance Transport Model follows the derivation by Stratton (Stratton 1972) which is derived starting from the Boltzmann Transport Equation. By applying certain assumptions, this model decomposes into the hydrodynamic model (Yang et al. 2003). Boltzmann statistics are used in preference to Fermi statistics. This model is modeled by following equations:

$$\xi_n = \frac{d(\ln \mu_n)}{d(\ln T_n)} = \frac{T_n}{\mu_n} \tag{1}$$

$$\xi_p = \frac{d(\ln \mu_p)}{d(\ln T_p)} = \frac{T_p}{\mu_p} \tag{2}$$

The parameters ξ_n and ξ_p are dependent on the carrier temperatures. Different assumptions concerning ξ_n and ξ_p correspond to different non-local models. In the high-field saturated velocity limit that corresponds to velocity saturation the carrier mobilities are inversely proportional to the carrier temperatures. We put $\xi_n = 0$, this would correspond to the simplified Hydrodynamic Model for electrons. The parameter, ξ_n can be specified using the KSN parameters on the MODELS statement (Atlas User's Manual 2014).

For the study of gate currents, the Selberherr's impact ionization and hot-electron-injection models are used to provide an accurate measure of hot-carrier-injection in short-channel devices. All the simulations have been performed using physical models accounting for the electric field-dependent and concentration-dependent carrier mobilities, Shockley–Read–Hall recombination/generation with doping dependent carrier lifetime. The mobility model used is the inversion layer Lombardi CVT mobility model, wherein concentration-dependent mobility, high field saturation model, and mobility degradation at interfaces are all included (Atlas User's Manual 2014). Default simulator coefficients for all parameters have been employed. In order to fairly analyze the device performances, all the three devices are optimized to have the same threshold voltage, i.e., 0.32.

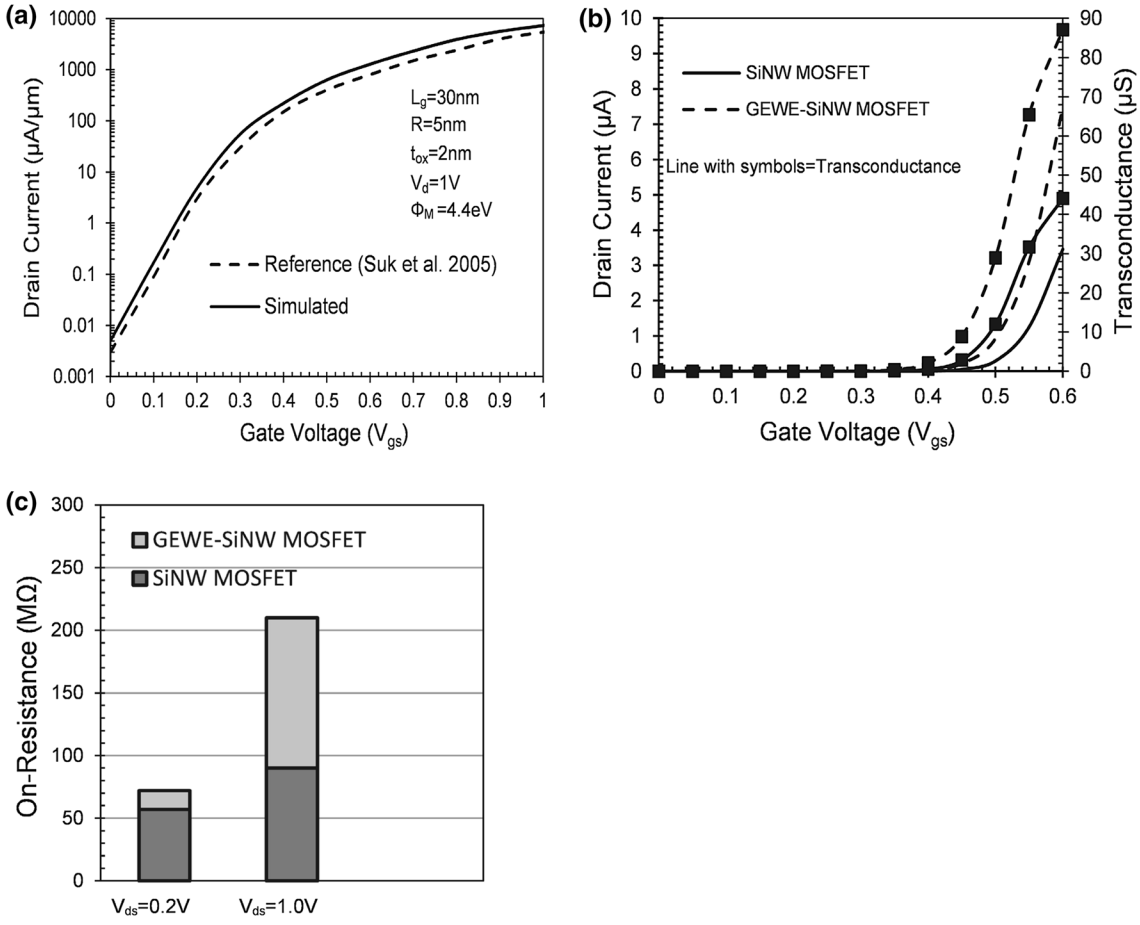


Fig. 2 **a** Calibration with the experimental result, $I_{ds} - V_{gs}$ characteristics of silicon nanowire MOSFET, **b** transfer and transconductance characteristics of silicon nanowire and GEWE-silicon nanowire MOSFET at $V_{DS} = 0.4$ V respectively, **c** On-resistance of SiNW and GEWE-SiNW MOSFET

3 Results and discussion

In this paper, we compare SiNW and GEWE SiNW MOSFETs with the conventional MOSFET. The present analysis is carried out for a channel length, $L_g = 30$ nm, radius of NW, $R = 5$ nm and thickness of oxide, $t_{ox} = 1.5$ nm, substrate doping is p-type with concentration of 1×10^{16} cm^{-3}; source and drain are n-type with uniform doping profiles is 5×10^{19} cm^{-3}. The metal gate work function at the source end (Φ_{m1}) is 4.8 eV (Au) and metal gate work function at the drain end (Φ_{m2}) is 4.4 eV (Ti) in GEWE and metal gate workfunction is 4.8 eV in SiNW MOSFET.

Figure 2a shows simulation result compared with fabricated SiNW MOSFET reported in literature (Suk et al. 2005). The calibration of model parameters used in simulation has been performed according to the published results to validate our simulations results. Figure 2b shows the drain current and gate transconductance characteristics of GEWE-SiNW MOSFET at drain bias 0.4 V. As shown in Fig. 2b, the current driving capability of nano-scale GEWE-SiNW MOSFET enhanced by 100 % in comparison to SiNW MOSFET. This enhancement is due to incorporation of metal gate engineering scheme which causes gradual step change of surface potential at the interface of two metals and hence increases gate efficiency of device. In

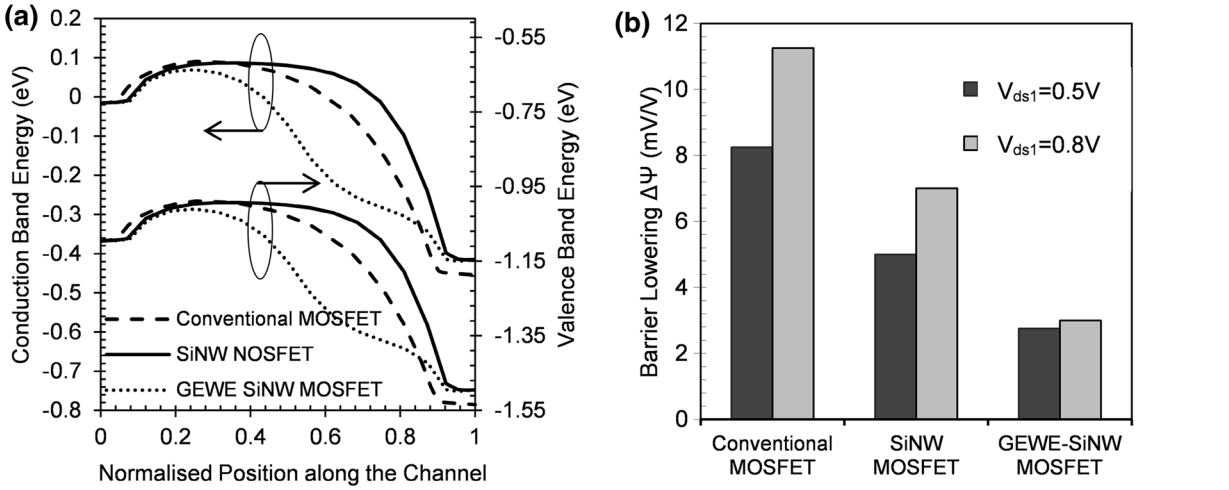


Fig. 3 **a** Conduction band energy and Valence band energy variation along the channel for $V_{gs} = 0.37$ V for conventional, SiNW and GEWE Si NW MOSFET at $V_{ds} = 0.55$ V, **b** DIBL versus V_{ds} variation for conventional MOSFET, SiNW MOSFET and GEWE Si NW MOSFET at $V_{gs} = 0$ V for two different values of V_{ds1}

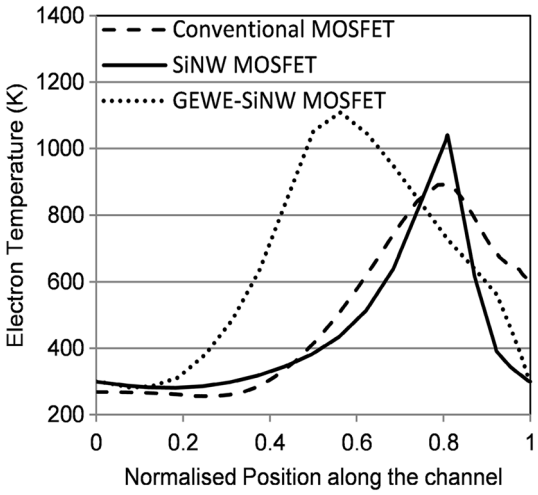


Fig. 4 Electron temperature variation along the channel for conventional MOSFET, SiNW MOSFET and GEWE Si NW MOSFET at $V_{ds} = 0.55$ V, $V_{gs} = 0.37$ V

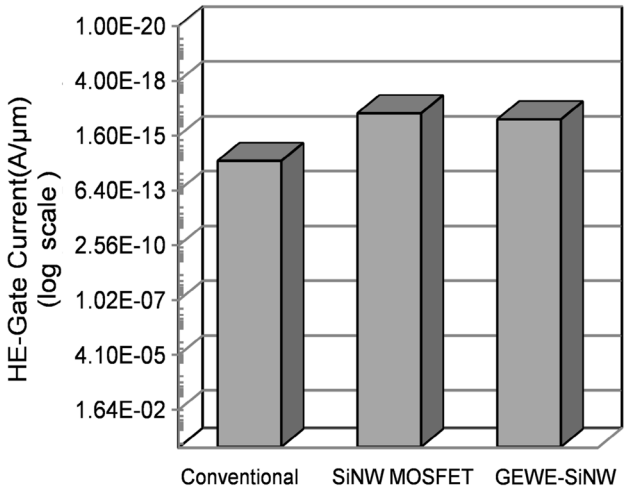


Fig. 5 HE gate current for conventional MOSFET, SiNW and GEWE Si nanowire MOSFET at $V_{ds} = 0.55$ V

general, the larger the transconductance figure for a device, the greater the gain (amplification) it is capable of delivering, when all other factors are held constant. It is clearly evident from Fig. 2b that transconductance of GEWE-SiNW is higher in comparison to conventional SiNW MOSFET due to enhanced on-current. As we know CMOS is used as a switch in linear region, and in saturation region on-resistance should be large so that we get large gain for amplifier Fig. 2c clearly shows that in linear and saturation region GEWE-SiNW MOSFET exhibit smaller and higher on-resistance respectively compared to SiNW MOSFET.

Figure 3a shows the variation of conduction band and valence band along the channel length for conventional MOSFET, SiNW and GEWE Si NW MOSFET at $V_{gs} = 0.37$ V and $V_{ds} = 0.55$ V. As shown in Fig. 3a, the significant barrier lowering just below the gate in conventional MOSFET leads to leakage current. The improvement in DIBL effect in GEWE is due to the step in conduction band energy profile in the channel region, which is resulting from metal gate workfunction difference. In SiNW and GEWE-SiNW, there is improved gate controllability over the channel due to cylindrical gate design hence DIBL is significantly lowered as shown in Fig. 3b in comparison to conventional MOSFET. $V_{ds} = 0.1$ V is used as reference for calculating barrier lowering in all three device. Furthermore, with an increase in the drain to source bias, the DIBL effect increases, but this increase is prominent in compared MOSFET comparison to the GEWE-SiNW. The DIBL is calculated as follows:

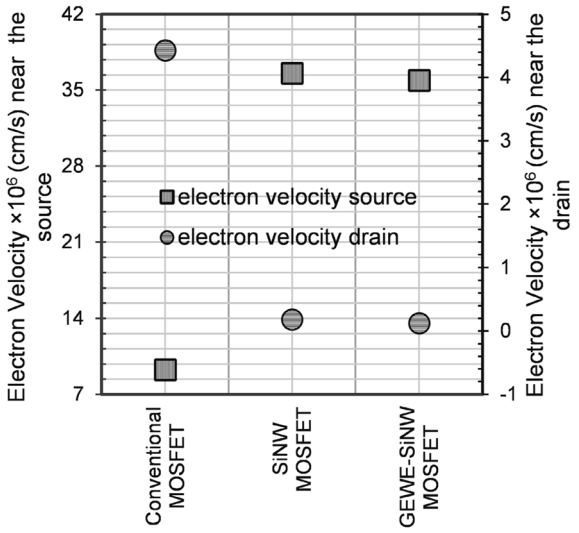


Fig. 6 Electron velocity for source and drain side for conventional MOSFET, SiNW and GEWE Si nanowire MOSFET at $V_{ds} = 0.55$ V, $V_{gs} = 0.37$

$$DIBL = \frac{V_{Th1} - V_{Th0}}{V_{ds1} - V_{ds0}} \tag{3}$$

where V_{Th1} is the threshold voltage of device when V_{ds1} bias is applied and V_{Th0} is the threshold voltage of device when reference bias voltage i.e. $V_{ds0} = 0.1$ V is applied.

Figure 4 predicts the variation of electron temperature along the channel length from source to drain for conventional, SiNW and GEWE-SiNW MOSFET. As it is clear from the results, an appreciable reduction in electron temperature for GEWE-SiNW at the drain side which is due to reduction in electric field at the drain side, which in turn reduces the leakage current and hence improves hot carrier immunity in comparison to conventional MOSFET. The reduction in electric field at drain side in GEWE-SiNW is because of gate electrode function engineering, which causes step potential profile at the interface of two metal gates (Long and Chin 1997).

Gate current of conventional MOSFET is comparatively higher than SiNW MOSFET because of short

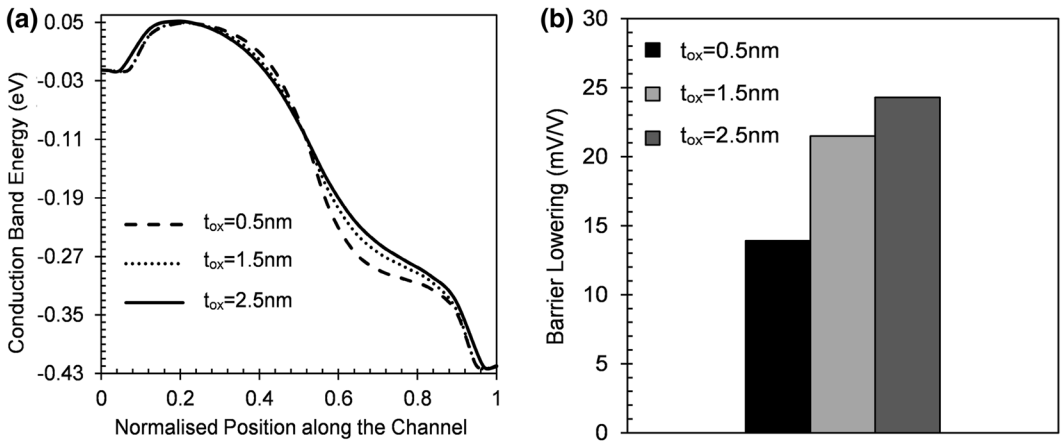


Fig. 7 **a** Conduction band profile for GEWE-SiNW MOSFET at different values of oxide thickness. **b** Barrier lowering for GEWE-SiNW MOSFET at different values of different values of oxide thickness

channel effects which exists in conventional MOSFET. Figure 5 shows the hot electron (HE) injected gate current of conventional, SiNW and GEWE-SiNW MOSFET at $V_{ds} = 0.55$ V. HE–I_g in SiNW MOSFET is reduced significantly in comparison to GEWE-SiNW MOSFET. Since, GEWE favors high electric field in the channel because of step potential profile in the channel, which causes tunneling of electrons from channel to gate oxide causing gate tunneling current thereby increasing HE-I_g current in comparison to SiNW MOSFET.

Figure 6 predicts the electron velocity of conventional, SiNW and GEWE-SiNW MOSFET at $V_{ds} = 0.55$ V along the channel. In order to improve the carrier transport efficiency in the channel, the carrier velocity near the source should be more than near the drain. Higher electron velocity (3.59×10^7 cm/s) near the source side is obtained for GEWE-SiNW in contrast to 3.5×10^7 and 0.9×10^7 cm/s obtained for SiNW and conventional MOSFET, respectively, as shown in Fig. 5, as a result enhancing the source carrier injection into the channel and hence results in hot carrier immunity in GEWE-SiNW MOSFET. For low-power and low-voltage analog circuit applications, lower electron velocity near the drain end are important design parameters.

3.1 Effect of gate oxide thickness

Figure 7a shows the variation of conduction band energy of GEWE-SiNW MOSFET with the normalized position along the channel. The result shows that as we scale down the oxide thickness, step in conduction band profile increases and hence results in lowering of DIBL as is indicated in Fig. 7b. It is clearly evident from the figure that when oxide thickness is 0.5 nm barrier lowering is 13.9 mV/V compared to 24.3 mV/V at 2.5 nm, thereby reflecting significant improvement in SCEs at sub-nm range. Figure 8 depicts the variation of electric field of GEWE-SiNW MOSFET at different oxide thickness with the normalized position along the channel. As is evident from the figure, the drain electric field reduces as oxide thickness decreases due to incorporation of lower work-function of metal gate and thinning of gate oxide, results in less number of hot-carriers at the drain end and thus improves hot-carrier reliability of device. Also, as shown in figure the peak electric field in the channel increases due to difference in metal gate workfunction, which results in

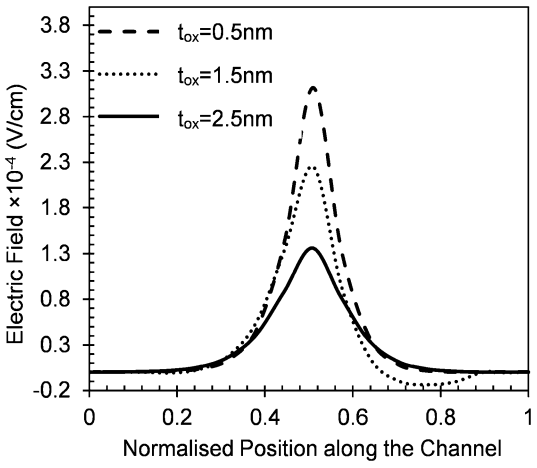


Fig. 8 Electric Field for GEWE-SiNW MOSFET at different values of oxide thickness

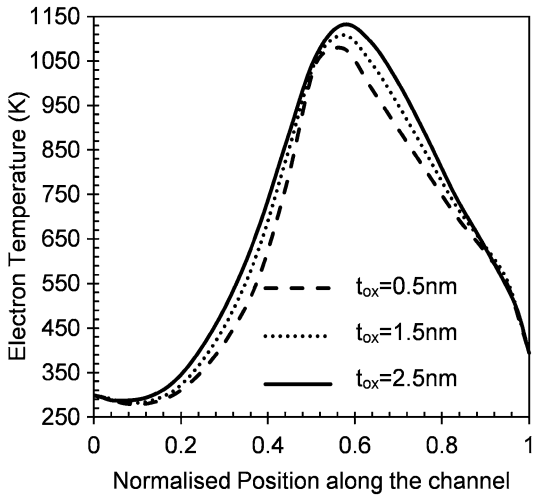


Fig. 9 Electron Temperature for GEWE-SiNW MOSFET for $V_{gs} = 0.6$ V and $V_{ds} = 0.4$ at different values of oxide thickness

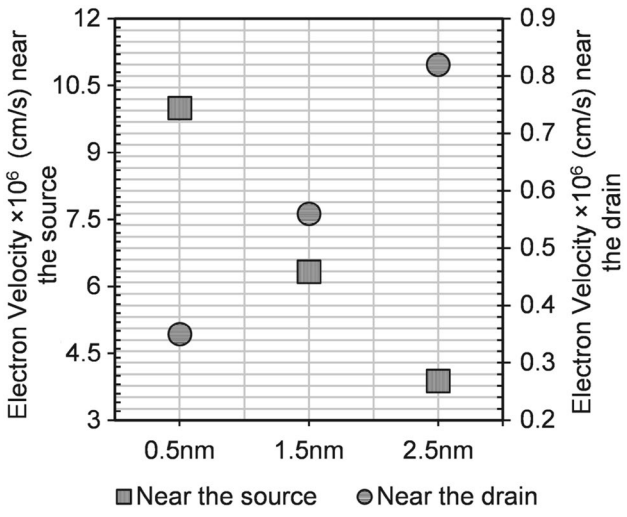


Fig. 10 Electron velocity for source and drain side for GEWE-SiNW MOSFET for $V_{gs} = 0.6$ V and $V_{ds} = 0.4$ at different values of oxide thickness

improvement in electron velocity and thus enhances the carrier efficiency.

Figure 8 depicts an appreciable reduction in electron temperature of GEWE-SiNW MOSFET at the drain end as we scale down the oxide thickness. This reduction in temperature is due to reduction in electric field at the drain end which results less number of hot-carriers and thus lowers the temperature, which accounts improvement in hot-carrier reliability. Figure 9 depicts an appreciable reduction in electron temperature of GEWE-SiNW MOSFET at the drain end as we scale down the oxide thickness. This reduction in temperature is due to reduction in electric field at the drain end which results less number of hot-carriers and thus lowers the temperature, which accounts improvement in hot-carrier reliability.

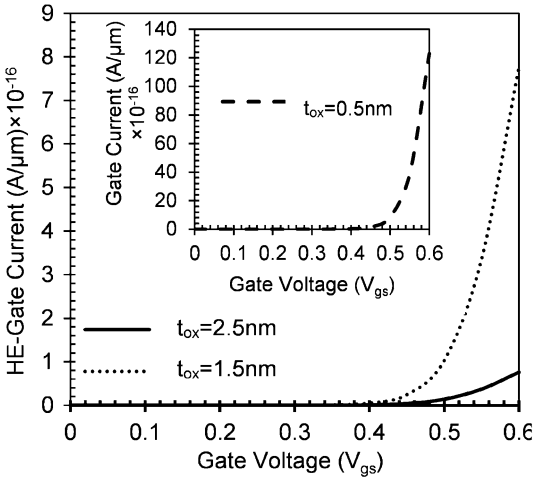


Fig. 11 Hot-electron injection gate current for GEWE-SiNW MOSFET at different values of oxide thickness

Figure 10 shows the effect of oxide thickness on electron velocity of GEWE-SiNW MOSFET. In order to upgrade the carrier efficiency, electron velocity near the source end should be greater than velocity at the drain end. Result depicted electron velocity at source side is 0.63×10^7 cm/s for 1.5 nm oxide thickness in contrast to 0.3×10^7 cm/s for 2.5 nm, thereby enhancing the carrier injection into the channel. If we further scale down the oxide thickness to 0.5 nm results in improvement in electron velocity near the source (1.2×10^7 cm/s), due to improved gate controllability over the channel and effectiveness of dual material gate, hence leading to improvement in carrier efficiency and driving current capability of device. Further, Fig. 10 also validates a lower electron velocity of 0.35×10^7 cm/s at the drain end for 0.5 nm thickness in contrast to 0.56×10^7 and 0.82×10^7 cm/s for 1.5 and 2.5 nm respectively, thus imparting outstanding hot-carrier immunity.

Hot-Electron Injection Gate Current (HE-I_g) is a key parameter to evaluate the hot-carrier reliability of a device. Figure 11 depicts the variation of gate current as s function of gate to source voltage. As is evident from figure that when oxide thickness goes on decreasing from 2.5 to 0.5 nm gate current increases, due to tunneling of electrons through the gate oxides which causes increment in gate current and thus hinders the device performance.

3.2 Effect of dielectric constant (high-k)

As we have seen in previous section, with scaling of oxide thickness below 1.5 nm such an ultra-thin gate oxide will induce high direct tunneling current (leakage gate current) therefore, it is required for high dielectric constant (k) material to replace SiO_2 as a gate dielectric to reduce standby power dissipation.

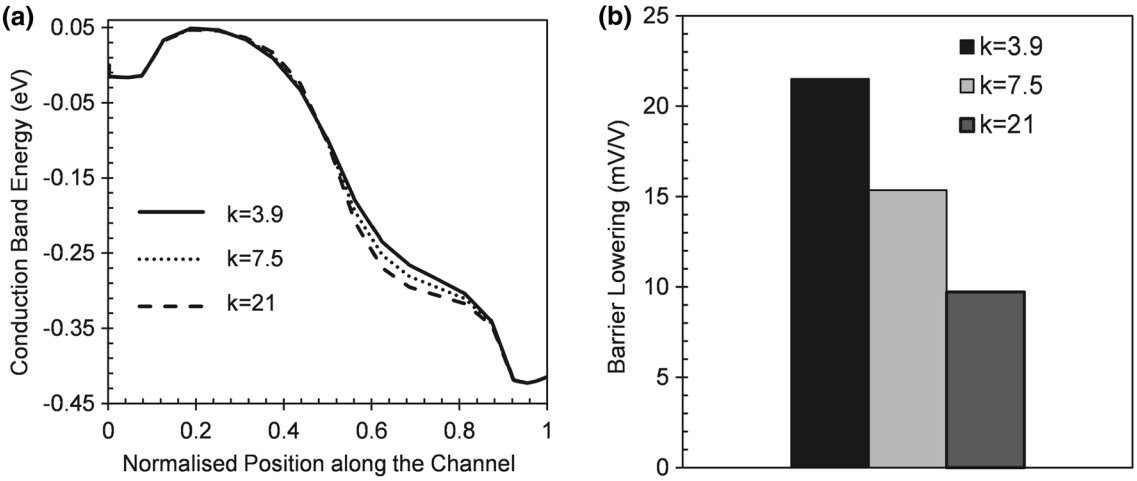


Fig. 12 **a** Conduction band and valence band energy variation along the channel for GEWE-SiNW MOSFET at different values of oxide thickness. **b** DIBL variation for GEWE-SiNW MOSFET for $V_{gs} = 0.6$ V and $V_{ds} = 0.4$ at different values of High-k

Figure 12a shows the variation of conduction band energy of GEWE-SiNW MOSFET at different values of high-k. As value of dielectric constant increases, the step in conduction band profile increases due to metal gate workfunction difference and also due to increased gate capacitance, thus providing improvement in DIBL. DIBL is significantly reduced as we increase the value of permittivity as is shown in Fig. 12b, thereby reflecting SCE immunity.

Figure 13 shows the variation of electric field with normalized position along the channel for GEWE-SiNW MOSFET. It is evident from the Fig. 13, the peak electric field at the interface of two metal gates increases with increase in permittivity value of insulators, which results in an improvement in average velocity of carriers. In addition, there is prominent reduction in field at the drain end due to incorporation of lower metal workfunction at the drain end and due to fact that gate capacitance is directly proportional to dielectric constant and as the dielectric constant increases the gate capacitance increases also the control of gate on the channel becomes stronger therefore enhances the driving capability of device and hence reduces field at the drain end.

Electron temperature in the channel decreases with increase in high k value as is evident from Fig. 14. As carriers move from source to drain, carrier temperature gradient is reduced significantly giving a remarkable reduction in electron temperature to 980 K for HfO_2 in comparison to 1018 K for SiO_2 as shown in Fig. 14, due to increased gate controllability of device and reduction of field at the drain end which results in enhancement of hot-carrier reliability.

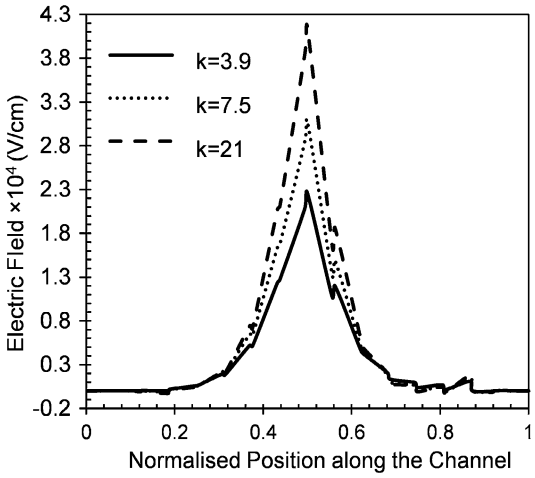


Fig. 13 Electric field for GEWE-SiNW MOSFET at different values of High-k

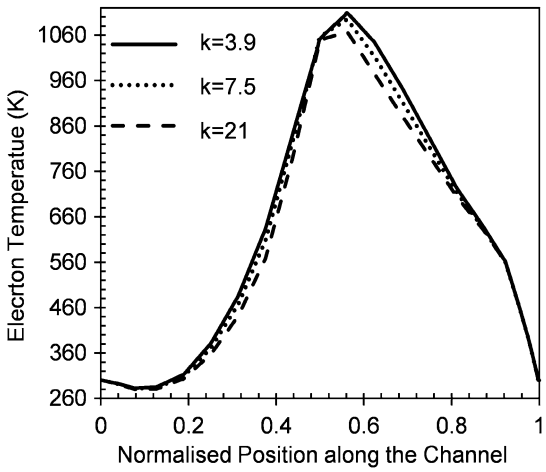


Fig. 14 Electron temperature variation along the channel for GEWE-SiNW MOSFET for $V_{gs} = 0.6$ V and $V_{ds} = 0.4$ at different values of High-k

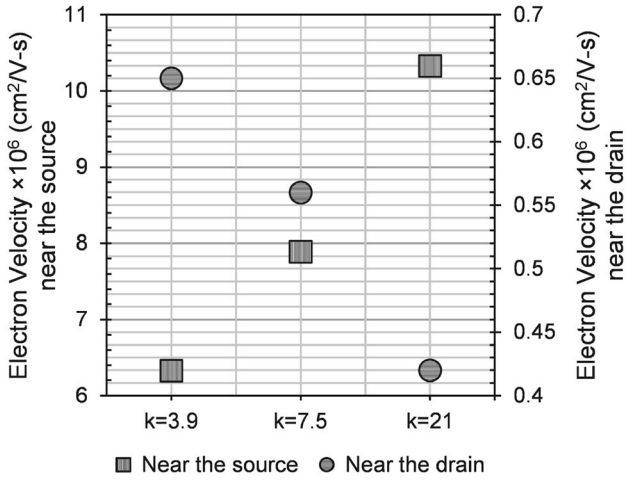


Fig. 15 Electron velocity for source and drain side for GEWE-SiNW MOSFET for $V_{gs} = 0.6$ V and $V_{ds} = 0.4$ at different values of High-k

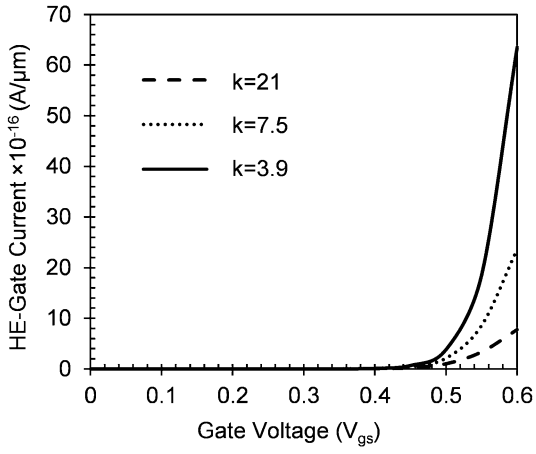


Fig. 16 Hot-electron injection gate current for GEWE-SiNW MOSFET at different values of High-k

Table 1 Extracted parameters at different values of oxide thickness and high-k

Permitivity value	Parameters		Oxide thickness		
			0.5 nm	1.5 nm	2.5 nm
k = 3.8 (SiO_2)	Electric field ($\times 10^4$ V/cm)	Near source	3.11	2.25	1.36
		Near drain	0.3	0.86	0.9
	Gate current ($\times 10^{-16}$ A/μm)		123	7.77	0.75
	Electron temperature (K)		1043	1108	1131
	Electron mobility (cm^2/V -s)		1092	1020	897
	Electron velocity ($\times 10^7$ cm/s)	Near source	0.99	0.634	0.38
		Near drain	0.035	0.056	0.082
k = 7.5 (Si_3N_4)	Electric field ($\times 10^4$ V/cm)	Near source	3.66	3.0	1.98
		Near drain	0.2	0.22	0.24
	Gate current ($\times 10^{-16}$ A/μm)		0.8	2.36	7.6
	Electron temperature (K)		1062.1	1091	1135.0
	Electron mobility (cm^2/V -s)		1112.1	1018	915.2
	Electron velocity ($\times 10^7$ cm/s)	Near source	1.2	0.78	0.56
		Near drain	0.013	0.056	0.004
k = 21 (HfO_2)	Electric field ($\times 10^4$ V/cm)	Near source	4.12	4.15	2.92
		Near drain	0.016	0.012	0.029
	Gate current ($\times 10^{-16}$ A/μm)		0.55	0.7	1.2
	Electron temperature (K)		1047	1065	1116.6
	Electron mobility (cm^2/V -s)		1127	1066	962
	Electron velocity ($\times 10^7$ cm/s)	Near source	1.33	1.55	0.82
		Near drain	0.0108	0.042	0.0105

Figure 15 demonstrates the electron velocity of GEWE-SiNW MOSFET at different values of dielectric constant. It is evident from result, as dielectric constant increases from 3.9 to 21 electron velocity at source increases and decreases at drain end due to enhancement in electric field in the channel thus improves the carrier efficiency and reliability of device at low power applications.

The hot-electron gate current is due to the injection of hot-carriers into the oxide and transport of these carriers towards the gate electrode. Figure 16 shows the HE-I_g of GEWE-SiNW MOSFET at $V_{ds} = 0.4$ V. It is clearly shown than with incorporation of HfO_2, gate current (7.77×10^{-16} A/μm) decreases appreciably in comparison to SiO_2 (6.35×10^{-15} A/μm). This reduction in gate current is due to large gate capacitance which inhibits the tunneling of electrons at 1.5 nm oxide thickness. Thus reducing the power dissipation and finds its efficacy in low power analog applications. Table 1 gives the hot-carrier reliability metrics values for three different oxide thickness and dielectric values.

4 Conclusion

The present work proposes GEWE-SiNW MOSFET for improved hot carrier reliability applications over conventional and SiNW MOSFET. The reduction in electron velocity and electron temperature near the drain end is achieved with GEWE-SiNW MOSFET which reflects its effectiveness in high-performance applications where device reliability is a major concern. Furthermore, notable reduction in SCEs such as DIBL in GEWE-SiNW increases its current drive ability and also there is significant reduction in hot-electron gate current in GEWE-SiNW MOSFET compared to conventional MOSET, which results in improving its standby power consumption in electrical appliances and logic gates. Further, the scaling of oxide thickness shows an outstanding performance by GEWE-SiNW MOSFET with an exception of high HE-I_g at 0.5 nm. With incorporation of high-k, HE-Ig reduces significantly and improves the device reliability at 300 K. Thus, an oxide thickness 0.5 nm and oxide dielectric 21 (HfO_2) with GEWE-SiNW MOSFET is an effective candidate for low power CMOS applications such as amplifiers and switching applications.

Acknowledgments The authors would like to thank the Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University (formerly DCE) and one of the authors (Neha Gupta) is grateful to the University Grant Commission (UGC) for providing the necessary financial assistance to carry out this research work.

References

Arora N (2012) MOSFET models for VLSI circuit simulation: theory and practice. Springer, New York

ATLAS User's Manual (2014) SILVACO Int., Santa Clara, CA

Barsan RM (1981) Analysis and modeling of dual-gate MOSFET's. IEEE Trans Electron Devices 28:523–534. doi:10.1109/T-ED.1981.20377

Chaujar R, Kaur R, Saxena M, Gupta M, Gupta RS (2008) TCAD assessment of gate electrode workfunction engineered recessed channel (GEWE-RC) MOSFET and its multilayered gate architecture—part I: hot-carrier-reliability evaluation. IEEE Trans Electron Devices 55:2602–2612. doi:10.1109/TED.2008.2003085

Chen X, Tan CM (2014) Modeling and analysis of gate-all-around silicon nanowire FET. Microelectron Reliab 54:1103–1108. doi:10.1016/j.microrel.2013.12.009

De V, Borkar S (1999) Technology and design challenges for low power and high performance. In Proceedings of international symposium on low power electronics and design 163–168

Deb S, Singh NB, Islam N, Sarkar SK (2012) Work function engineering with linearly graded binary metal alloy gate electrode for shortchannel SOI MOSFET. IEEE Trans Nanotechnol 11:472–478. doi:10.1109/TNANO.2011.2177669

Dubey S, Santra A, Saramekala G, Kumar M, Tiwari PK (2013) An analytical threshold voltage model for triple material cylindrical gate all around (TMCGAA) MOSFETs. IEEE Trans Nanotechnol 12:766–774. doi:10.1109/TNANO.2013.2273805

Groeseneken GV (2001) Hot carrier degradation and ESD in submicrometer CMOS technologies: how do they interact? IEEE Trans Device Mater Reliab 1:23–32. doi:10.1109/7298.946457

Gupta N, Chaujar R (2014) Implications of transport models on the analog performance of gate electrode work function engineered (GEWE) SiNW MOSFET. IEEE 2nd international conference on devices, circuits and systems (ICDCS), 1–5. doi:10.1109/ICDCSyst.2014.6926154

Iwai H, Natori K, Shiraishi K, Iwata J, Oshiyama A, Yamada K, Ohmori K, Kakushima K, Ahmet P (2011) Si nanowire FET and its modeling. Sci China Inf Sci Springer 54:1004–1011. doi:10.1007/s11432-011-4220-0

Jeon HJ, Kim YB, Choi M (2009) Standby leakage power reduction technique for nanoscale CMOS VLSI systems. IEEE Trans Instrum Meas 59(5):1127–1133. doi:10.1109/TIM.2010.2044710

Kaur R, Chaujar R, Saxena M, Gupta RS (2007) Hot-carrier reliability and analog performance investigation of DMG-ISEGaS MOSFET. IEEE Trans Electron Devices 54:2556–2560. doi:10.1109/TED.2007.902855

Kranti A, Armstrong GA (2006) Engineering source/drain extension regions in nanoscale double gate (DG) SOI MOSFETs: analytical model and design considerations. Solid State Electron 50:437–447. doi:10.1016/j.sse.2006.02.012

Krieger G, Sikora R, Cuevas PP, Misheloff MN (2001) Moderately doped NMOS (M-LDD)-hot electron and current drive optimization. IEEE Trans Electron Devices 38:121–127. doi:10.1109/16.65745

Long W, Chin KK (1997) Dual material gate field effect transistor (DMGFET). Electron devices meeting tech dig 549–552. doi:10.1109/IEDM.1997.650445

Momose HS, Ono M, Yoshitomi T, Ohguro T, Nakamura S, Saito M, Iwai H (1996) 1.5 nm direct-tunneling gate oxide Si MOSFET's. IEEE Trans Electron Devices 43:1233–1242. doi:10.1109/16.506774

Mukherjee C, Maiti CK (2012) Silicon nanowire FinFETs. In: X. Peng (ed) Nanowires: recent advances. doi: 10.5772/52591

Nirmal D, Vijayakumar P, Thomas DM, Jebalin BK, Mohankumar N (2013) Subthreshold performance of gate engineered FinFET devices and circuit with high-k dielectrics. Microelectron Reliab 53:499–504. doi:10.1016/j.microrel.2012.09.008

Orouji AA, Kumar MJ (2005) Shielded channel double-gate MOSFET: a novel device for reliable nanoscale CMOS applications. IEEE Trans Device Mater Reliab 5:509–514. doi:10.1109/TDMR.2005.853505

Pan Y, Ng KK, Wei CC (1994) Hot-carrier induced electron mobility and series resistance degradation in LDD NMOSFE's. IEEE Electron Device Lett 15:499–501. doi:10.1109/55.338416

Sharma S (2012) Comparative analysis of low power and high performance PTM models of CMOS with HiK-metal gate technology at 22 nm. Int J Soft Comput Eng 2:257–261

Sharma D, Vishvakarma SK (2013) Precise analytical model for short channel cylindrical gate (CylG) gate-all-around (GAA) MOSFET. Solid-State Electron 86:68–74. doi:10.1016/j.sse.2012.08.006

Sinha SK, Chaudhury S (2013) Impact of oxide thickness on gate capacitance—a comprehensive analysis on MOSFET, nanowire FET, and CNTFET devices. IEEE Trans Nanotechnol 12:958–964. doi:10.1109/TNANO.2013.2278021

Sohn CW, Kang CY et al (2012) Device design guidelines for nanoscale FinFETs in RF/analog applications. IEEE Electron Device Lett 33:1234–1236. doi:10.1109/LED.2012.2204853

Stratton R (1972) Semiconductor current-flow equations (diffusion and degeneracy). IEEE Trans Electron Devices 19:1288–1292. doi:10.1109/T-ED.1972.17592

Suk SD, Lee S-Y, Kim S-M et al (2005) High performance 5 nm radius twin silicon nanowire MOSFET(TSNWFET): fabrication on bulk Si wafer, characteristics, and reliability. IEEE IEDM Tech Digest. doi:10.1109/IEDM.2005.1609453

Venkatesan S, Neudeck GW, Pierret RF (1992) Dual-gate operation and volume inversion in n-channel SOIMOSFET's. IEEE Electron Device Lett 13:44–46. doi:10.1109/55.144946

Wang R, Zhuge J, Huang R, Tian Y, Xiao H, Zhang L, Li C, Zhang X, Wang Y (2007) Analog/RF performance of Si nanowire MOSFETs and the impact of process variation. IEEE Trans Electron Devices 54:1288–1294. doi:10.1109/TED.2007.896598

Wang T, Lou L, Lee C (2013) A junctionless gate-all-around silicon nanowire FET of high linearity and its potential applications. IEEE Electron Device Lett 34:478–480. doi:10.1109/LED.2013.2244056

Yang G, Wang S, Wang R (2003) An efficient preconditioning technique for numerical simulation of hydrodynamic model semiconductor devices. Int J Numer Simul Electron Netw Devices Fields 16:387–400. doi:10.1002/jnm.507

Yu B, Wann CHJ, Nowak ED, Noda K, Hu C (1997) Short-channel effect improved by lateral channel engineering in deep-submicronmeter MOSFET's. IEEE Trans Electron Devices 44:627–634. doi:10.1109/16.563368

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), Volume 5, Special Issue 1, April 2015)
Second International Conference on Advanced Developments in Engineering and Technology (ICADET-15), INDIA.

Promoting Human Resources Skill Developments and Impowering India with Solar Power Technologies, Education, Training and Innovation

Dr. L P Singh[1], Dr. J P Kesari[2], Dr. A K Ojha[3], Mr. Gaurav Kumar Gaur[4]
[1]*Asst. Professor, SHIATS, Allahabad, India*
[2]*Associate Professor in DTU, Delhi, India*
[3]*Associate Professor, 4Asst. Professor, SIET, Hapur, India*
[1]lpsingh_76@yahoo.co.in, [2]drjpkesari@gmail.com, [3]ashokojh@gmail.com, [4]gauravvviet@gmail.com

I. Introduction

India has tremendous energy needs for its vast growing population but is finding it very difficult to meet those needs through existing conventional means of power generation. Electricity consumption in India has been increasing at one of the fastest rates in the world due to population growth and rapid economic development. India's economic growth has been some what stalled because the amount of energy generated is inadequate to keep the growth engine of India running at the desired speedand there are huge energy shortages almost in every states of the country.

Now, The question arises,how can India solve the existing problem to meet the future energy demands and help eliminate wide –ranging power outages in the future? To solve this question, answer is, "India is blessed with plentiful of renewable energy sources and if those are properly utilized, India can solidify it's place in the world as a super power which can not only take care of it's own energy needs but also of other south east Asian nations.

Renewable energy also has the advantage of allowing decentralized generation and distribution of energy — particularly for the rural energy needs, and thereby empowering people at the grass root level.

II. Indian Rural Scenario:

Majority of India's population still lives in villages. The population in villages has grown substantially in the past fifty years. This growth has put tremendous pressure on existing resources needed to support them.Census of India 2011 indicates that access to electricity stands at only around 55.5% of the rural households (i.e. 92.8 million household out of 167.8 million rural households). In few states the household electrification level is strikingly low (Bihar -10.4%, Assam 28.4%, Uttar Pradesh 23.8%, Orrisa 35.6, Jharkhand 32.3%, and West Bengal 40.3%).

Firewood, crop residue and cow dung cake continues to be the most widely used energy sources in India , with around 86% of the rural household dependent upon it. In addition 23.2% of urban household still relies in traditional fields to meets their energy needs. This has resulted in enormous pollution of air, water and land.Clearly, skilled development andinnovations for clean energy technologies are needed to accommodate the growing population needs in India for healthier environment. The abundant presence of solar insolation at any location in a village make its uses much attractive for creating such a type of environment.

The examples of clean energy technologies that are in wide scale use in rural India include PV powered mobile phones and network towers, electrical appliances like Tv, vaccine refrigerators, fans, motors and pumps , farm equipments, water desalinator etc.The cost and reliability of these energy sources has now been establishedthroughout India.

III. Solar Energy Potential And Technologies:

Solar is the prime free source of inexhaustible energy available all over the globe and Indiais endowed with vast solar energy potential. About 5,000 trillion kWh per yearsolar energy is incident over Indian mass land with most parts receiving insolation of 4-7 kWh per sq. m per day. Even if a tenth of this potential is utilized, it could mark the solution for India's power problems — by using the country's deserts and barren lands for construction of solar plants. Renewable energy has the potential to re-energize India's economy by creating millions of new jobs, allowing the country to achieve energy independence, reduce its trade deficits and propel it forward as a "Green Nation". In short, renewable energy offers large pool of benefits for India to propel its development.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), Volume 5, Special Issue 1, April 2015)
Second International Conference on Advanced Developments in Engineering and Technology (ICADET-15), INDIA.

India should take full advantage of this golden opportunity because renewable energy has particular relevance in remote and rural areas, where there are around 500 million people who don't have access to reliable sources of energy.

➢ *Solar PV stand alone systems:*

The concept of Solar Photovoltaic power plants can be used as decentralized power plant to power entire villages.In this case, the solar energy is harvested and stored and inhabitants can charge devices from energy stored in batteries. 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.

➢ *Solar PV Powered Refrigeration Facility:*

Farm products degrade in normal climatic conditions.Lowtemperatures
and less humidity allows these products to have extendable their usable life. Food preservation will greatly help to satisfy India's growing demands.Storages with compartmental design can be built in rural areas and their environment can be regulated with help of solar energy technology. These storages can be built through village cooperatives and then rented to farmers for their use.

➢ *Solar Powered Agricultural pumps:*

In a typical village in rural India, the villagers have community water wells to draw water from underground.They also have water wells for irrigating fields. Most of these are operated manually and those that operate on electric power are limited by the availability of such power. The power availability in rural India for use by farmers is extremely limited. Solar powered water wells will undoubtedly increase crop yield and better the quality of life of the inhabitants.

➢ *Centralized Solar street lighting system:*

Centralized Solar street lighting system can be used in many villages to provide safety and needed light to enable an individual to accomplish jobs in dark nights. Centralized PV system is installed at any village rooftop and pool of batteries are used to store energy generated during daytime. This energy is used to run street lights in night.

➢ *Solar Powered Farm Implements:*

Farm Tools are essential for farmers. From digging to harvesting the crops, several implements /tools are needed.

Many of these are manually operated presently to save energy use. Solar powered farming will realize great leap forward in productivity and quality of life of farmers.

➢ *Solar Powered Vehicles:*

For fast transportation of materials, crops, animals, and humans, rural areas now depend on vehicles that run on gas/petrol. These vehicles could be designed to run on solar power by introducing plug in hybrid solar vehicles.

➢ *Solar Powered Waste Management:*

Farms and farm homes produce waste which largely pollutes land, water, and air therefore becoming a health hazard, in addition to being an eye sore. Solar powered waste recyclesystems can be useful products would benefit immensely the rural communities.

➢ *Solar Desalinatior and Purification:*

Solar Sills can be used at wide scale to desalinate saline sea water into potable water in water deficient areas of semi arid India. Also commercially available Solar Electroculators are available to be used as community level water filtering systems.

➢ *Solar Powered Mobile network towers:*

Major telecom operators in India still uses huge amount of diesel to run their network towers in remote locations. PV systems with huge capacity can be used in synchronisation of diesel Generators to run mobile towers even in remote places in mountains. this will help telecom operators to offset their diesel consumption and make the total operation more profitable.

➢ *Solar thermal Mass cooking systems:*

Solar concentrators like Scheffler reflectors can be used in community run village schools and Anganwadis to prepare food for mid day meal scheme instead of using LPG or locally sourced biomass.

IV. Major Steps To Harness Renewable Energy

Government should Aggressively expand large-scale deployment of both centralized and distributed renewable energy including solar, wind, hydro, tidal,biomass, and geothermal to ease the strain on the present transmission and distribution system – and allow more off-grid populations to be connected to Micro grid.

Facilitate growth in large scale deployment by installing 100 GW capacity solar power generation through both centralized and distributed energy within the next 10 years.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), Volume 5, Special Issue 1, April 2015)
Second International Conference on Advanced Developments in Engineering and Technology (ICADET-15), INDIA.

Form and implement strictly National Renewable Energy Policy of 30 percent penetration of total installed capacity by 2022 — to create demand, new industries and innovation, and a new wave of green jobs.Develop favourable Government policies to ease the project permitting process, and to provide start-up capital to promote the exponential growth of renewable energy. Establish national renewable infrastructure bank for funding renewable energy projects at low interest rate.Accelerate local demand for renewable energy by providing preferential Feed-in-Tariffs (FIT), Net Metering schemes and other incentives such as accelerated depreciation; tax holidays; initiatives for international partnerships/collaboration incentives for new technologies; human resources development; zero import duty on capital equipment and raw materials and excise duty exemption.Establish R&D facilities within academia, research institutions, industry, Government and civil society to stream line pathway for technological development and achieving indigenesity in renewable energy sector.Accelerate the development via implementation of Solar and Wind farms; utility-scale solar and wind generation nationwide.

V. Renewable Energy Education And Training : Role Of Delhi Technological University

5.1 Objectives Of The Centre of Energy to be established at DTU:

Delhi technological university must establish a centre of excellence on renewable energy technology with following objectives :

- To provide education on Renewable Energy Technology
- To engage in Research & Development in Solar Photovoltaic and Concentrated Thermal Energy Technologies and their various applications for power generation, air conditioning, process heat in various industries.
- To promote Renewable Energy Innovation Park for general public, RWA, Women, children, NGOs etc.
- To carry out training program me(s) in the field of Renewable Energy Technology for students of other Engineering/Polytechnic/ ITIs/Arts, Science and Commerce colleges.
- To Collaborate with Industries and national and international agencies related to promotion of education and research.

5.2Major Activities of the Proposed Centre

5.2.1 Higher Education:

- B. Tech. (Energy Technology)
- M. Tech. (Renewable Energy Technology0
- PhD Program me with specialization in Solar thermal and photovoltaic energy technology.

5.2.2Research & Development In The Proposed Centre:

5.2.2.1 Solar Photovoltaic

- Use of solar photovoltaic's in the development of solar powered passenger car
- Fundamental research in PV-related materials and cost reduction;
- Development of PV cells from several material

5.2.2.2 Solar Flat/Concentrated Thermal Energy system

- Cross Linear Concentrated Solar Power technology for solar process heat, refrigeration and electrical power generation.
- Solar process heat and space cooling.
- Concentrating solar systems for steam generation for cooking.
- Development of new designs and manufacturing processes for solar components and systems with an emphasis on improved performance, reliability and service life.

5.2.2.3Education and Training Programmes

- The Centre will conduct specialized training programmes for Participants from Industries, RET users , Students and faculties from Polytechnics/ Engineering colleges/ Universities

5.2.2.4 Electric Vehicle Modeling, Wireless Charging Systems:

- Use of solar photovoltaic in the development of electric passenger car
- Fundamental research in PV-related materials for wireless battery charging and cost reduction

5.2.2.5 Bio Energy Technologies:

- Carbon capture and sequestration.
- Bio Diesel production using MERS Technology
- Biomass densified briquettes and pallets for use in boiler applications
- Bio Gasifier and digester related Technologies
- Methane Production from Sanitary Land fill for power generation

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459 (Online), Volume 5, Special Issue 1, April 2015)
Second International Conference on Advanced Developments in Engineering and Technology (ICADET-15), INDIA.

- Biogas Generation from Kitchen, vegetables and Animal waste

5.2.3 Renewable Energy Innovation Park

- This will be the front face of the proposed Centre and showcase the latest trends & technologies in Renewable energy area of relevance to the world.
- Solar Water Heating Applications
- Solar Air Heating Systems Applications.
- Large Scale grain Dryers.
- Residential and Office Space Heating.
- Fruit and Vegetable Drying for off seasons Applications.
- Solar Steam Cooking in DTU's Boys and Girls Hostels.

5.2.4 Knowledge Partners of the proposed Centre of Excellence

- The Centre shall work in close cooperation with MNRE &National Institute of Solar Energy and shall collaborate with world class Universities and Institutions in India & abroad. Special features of the collaborative program shall be focus on development of new and efficient solar technologies & collaborative research in new horizons of Renewable Energy.

DTU should collaborate with major IIT's, NIT's and other Institutes of higher learning's and demand funds from AICTE, and other bodies such as Dep't of Environment & Forest of GNCT Delhi, Ministry of Environment & Forest (MoEF), GOI, UGC, DST, and International organizations such as UNESCO, UNDP, IREDA, SDA, GIZ, IDRC, SIDBI, etc

(a) Indian Knowledge Partner

Indian Universities & R&D Institutions
NISE, MNRE
Bergen Group
Sun borneEnergy Systems
Thermax limited
Rajiv Gandhi Technological University, Bhopal
National Thermal Power Corporation
Maharishi Solar Technologies Pvt Ltd

(b) International Knowledge Partners

Tokyo Institute of Technology, Japan
Asia Pacific Sunbelt Development Association, Japan
Toyo Engineering, Tokyo, Japan,
Ricoh, Japan
University of South Florida

VI. CSP Progress In India

The JNNSM mission document clearly specifies that all applications – domestic or industrial should be solarised. It also set a target of achieving deployment of 20 Million Square meters of solar collectors by the end of the Mission in 2022. The Solar Thermal Collectors include Solar Water Heaters, Solar Cooking, Solar Cooling, industrial process heat applications, among others. When it comes to grid-connected CSP projects, a total of 470 MW was allocated during the Batch 1 of the Phase 1 of the JNNSM. These projects were supposed to be commissioned by March 2013, but only 50 MW of the total 470 MW capacity has been commissioned so far. The delay has been attributed to lack of accurate Direct Normal Incidence (DNI) data, dusty conditions where the projects are located, technology related issues and supply shortage of the Heat Transfer Fluid (HTF) used in the projects.Theseprojectsare being implemented by MNRE and consortium partners consisting of Tata Power, Tata Consulting Engineers, Larsen & Toubro, Clique developments, KIE Solartherm.

Concentrated Solar Thermal Energy Technology based on Scheffler Dishes collectors

Scheffler concentrator is a fixed focus elliptical dish comprising of large number of mirrors to reflect sun rays to fixed receiver which contains working fluid to be heated. Working fluid is usually water which can be used as pressurized hot water or converted to steam. Suitable thermic fluid can also be used. Dish which automatically tracks the sun in east west direction from morning to evening is single tracked system. For the seasonal variation in sun position in north south direction, adjustments have to be made manually for every few days by operator. Scheffler dish was initially developed for direct cooking by concentrating solar energy in kitchen but now the dishes are also used for steam generation for process heat.

VII. Mitigation Of Climate Change And Low Carbon Initiatives

Impactsof Climate Change are already being faced by people all around the globe. However, every individual needs to take necessary measures in cutting carbon emissions through awareness campaigns that help in bringing out a change in attitude and consumption patterns of masses. Sustainable use of natural resources should be demonstrated in the awareness campaigns.

It is a very important to preserve the traditional knowledge and practices for providing effective sustainable life-solutions. Some of the initiatives at Institutional and Individual level are as follows: use of solar energyfor cooking, lighting, air conditioning, power generation etc, replacinghome appliances with the energy efficient one, switchinglighting systems to CFLs/LEDs, switch offthe lights and other electrical gadgets when not in use, temperature of AC'sto be maintainedabove 24 degree C, usewarm clothes instead of room heaters in the winter and much more.

VIII. Concluding Remarks

India has now emerged as a global power in this era. With anominal GDP growth rate of 7%, it is the world's fourth largest country in terms of purchasing power, the second most populous country in the world and a leading player in IT, telecom and business outsourcing. The government of India recognizes that in order to sustain this growth rate and strengthen its national competitiveness, solar energy innovation has a critical role to play in India. The Government of India's Solar Mission is a visionary and inspiring policy measure that has the potential to be a leading example for rest of the world.

Harnessing the potential of solar energy for the benefit of rural areas of India will improve the entire lifestyle of the rural population. Aggressive Research on all aspects of solar energy including science, technology, engineering, economics, and management will engage wide group of faculty and students as a team.

This can energize the entire University for a worthy cause. India's population lives in rural areas and the quality of life will positively be affected by innovative use of solar energy.

REFERENCES

[1] KESARI J.P., P.B. Sharma(2012) Empowering India with Solar Energy Technology: A Roadmap at International Conference on Climate Change and Sustainable management of natural Resources Feb 5th to 7th, 2012 at ITM University Gwalior.

[2] KESARI J.P.,KUMAR KRISHEN AND P B SHARMA (2012) Solar Power is a Panacea for Changing Rural India; The role of CSP under Indo-Japan Collaboration, Proceedings of 2nd International Symposium –collaborative research project of CSP for India and Japan , 25 April 12, Tokyo Institute of Technology, Tokyo, Japan.

[3] KESARI J.P., (2012) Eco-spirituality for sustainable living – Role of Solar Energy Technology', SSTech2012 National Conference at Mumbai University Oct 5th and 6th 2012.

[4] KESARI J.P., P.B. Sharma (2012) Developing Solar Energy Projects: A Doorstep Opportunies for India. Paper presented at Clean Energy and energy conservation 2012, organised by The Institution of Engineers(INDIA) 13-14 Oct ,2012 Nagpur.

[5] Kesari J P and P.B. Sharma (2013) Changing the face of India with Solar Power Technologies: Generation of Energy, Employment and Entrepreneurship. Paper presented at National Seminar on IT Applications in Energy Management, Rajiv Gandhi Institute of Information Technology, Amethi, A campus of IIIT Allahabad, April 15, 2013.

[6] KESARI J P and P B SHARMA(2013) Changing the Face of India with Solar Power Technologies: Generation of Energy, Employment and Entrepreneurship , Paper published in GRIDTECH 2013, April 3-5, PragatiMaidan, New Delhi.

Geotech Geol Eng
DOI 10.1007/s10706-015-9896-2

ORIGINAL PAPER

Shear Behaviour of Rock Joints Under CNL and CNS Boundary Conditions

Amit Kumar Shrivastava · K. Seshagiri Rao

Received: 22 October 2014 / Accepted: 16 May 2015
© Springer International Publishing Switzerland 2015

Abstract Shear behaviour of rock joints can be studied under both constant normal load (CNL) and constant normal stiffness (CNS) boundary condition. CNS condition is suitable for non planar and reinforced rock joints whereas CNL condition is suitable for planar and non reinforced rock joints. In the present study shear behaviour of modelled rock joints with different asperity have been experimentally investigated under both CNL and CNS boundary conditions. Test results indicate that CNS boundary conditions gives higher shear strength as compare to CNL boundary condition when other parameter of testing is kept same. But, this effect tends to diminishes with increase in normal stress on the shearing plane and at high normal stress both CNL and CNS boundary conditions gives same shear strength. A new shear strength model is proposed for both the boundary condition and the proposed model is validated by comparing the predicted shear strength with present experimental results and results available in the literature for natural and artificial rock joints with different asperity and roughness. The new study and model will be useful for safe and economical design of underground openings in jointed rocks, stability analysis of rock slopes, design of foundation on rock and design of rock socketed piles.

Keywords Shear strength · Dilation · Constant normal load · Constant normal stiffness · Modelling · Physical modelling

1 Introduction

Rock is heterogeneous and discontinuities are inevitable part of the rock masses. The presence of these discontinuities in the rock mass is in the form of joints, faults, bedding planes or other recurrent planar fractures. The presence of these discontinuities reduces the shear strength and influences shear behaviour. In addition to joints, shear behavior of the rock mass also influenced by various parameters like, joint roughness, stiffness of the surrounding rock mass, shear rate, condition of the joint i.e. unfilled/infilled, infill type and their thickness. The correct evaluation of shear behavior of rock mass is only possible if all these factors are considered during experimental, analytical and numerical studies. The correct evaluation of shear behavior of rock joints is important for safe and economical design of underground openings in jointed rocks, stability analysis of anchored/free rock slopes, risk

A. K. Shrivastava (✉)
Department of Civil Engineering, Delhi Technological University, Bawana Road, New Delhi 110042, India
e-mail: aksrivastava@dce.ac.in

K. S. Rao
Department of Civil Engineering, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110016, India
e-mail: raoks@civil.iitd.ac.in

Published online: 23 May 2015

assessment of underground waste disposal, design of foundation on rock and design of rock socketed piles (Rao et al. 2009).

The shear behavior can be investigated broadly under two different boundary conditions i.e. constant normal load (CNL) and constant normal stiffness (CNS). The shearing under CNL boundary conditions are suitable for situations: (1) planar joints, where no dilation takes place during the shearing process, (2) non planar and unreinforced rock slope where the surrounding rocks freely allow the joints to shear without restricting the dilation, thereby keeping normal load constant during shearing process. If the dilation of rock joints is either partly or fully restricted by the surrounding rock mass or reinforcement provided at the joints during the shearing process than there will be change in the normal load on the shearing plane. This change in the normal load on the shearing plane depends upon the stiffness of the surrounding rock mass or reinforcement provided and dilation resisted, hence shearing will take place under variable normal load condition, which can be simulated as CNS boundary conditions. Stiffness is the material property, which is constant for a given rock mass at a particular depth. As mostly rock joints are non planar and reinforcements are provided to the rock joints to increase the stability, hence a more representative behavior of joints would be achieved if shear behavior is investigated under CNS boundary conditions.

Conventional equipment used by different scholars (Newland and Allely 1957; Patton 1966; Goldestin et al. 1966; Ladanyi and Archambault 1970; Barton 1973; Byerlee 1975; Yang and Chiang 2000; Saiang et al. 2005; Ghazvinian et al. 2010) in previous days fails to explain correctly the shear behavior of rough and reinforced rock joints because of the limitation of boundary conditions, these equipment work under CNL boundary condition only. The shear strength model proposed by them based on these studies fails to predict shear strength correctly because of the limitation of boundary conditions.

To overcome this limitation, either conventional equipments were modified by Obert et al. (1976) and Ooi and Carter (1987) or new equipment developed by Indraratna et al. (1998) or servo controlled equipment developed by Jiang et al. (2004) and Kim et al. (2006) to conduct tests under CNS boundary conditions. The equipment modified or developed by Obert et al. (1976), Ooi and Carter (1987) and Indraratna et al. (1998) have difficulties to change the reaction rods, beams and stiffness plate to simulate different stiffness conditions according to the stiffness of the surrounding rock mass, while changing there is also a chance that sample may fail before testing of samples start. These difficulties have been overcome by using servo controlled equipment such as used by Jiang et al. (2004), but this equipment is suitable for comparatively smaller size of sample, due to that influence of joint roughness cannot be well studied and there is no facility to measure the rotation of the sample. The equipment developed by Kim et al. (2006) can test the samples only under two extreme conditions of stiffness of the boundary i.e. CNL (normal stiffness $k_n = 0$) and infinite normal stiffness condition ($k_n = \infty$). Servo control large scale direct shear testing machine designed by Shrivastava and Rao (2013) is capable of conducting tests on large and jointed prismatic specimens through friction free rigid platens under CNL and CNS boundary conditions, which can test under any condition of stiffness of the boundary without using or changing the reaction rods, beams and stiffness plates.

The present study will explain the effect of the joint roughness and boundary conditions on the shear behavior of the rock joints and suitable failure criteria will be proposed.

2 Testing Equipment

A servo control large scale direct shear apparatus designed and developed by Shrivastava and Rao (2013) is used for testing the rock joint under CNL and CNS boundary conditions. The apparatus consists of three main units such as loading unit, hydraulic power pack with servo valve and data acquisition and controlling unit as shown in Fig. 1. The loading unit consist of two shear box of size 300 mm × 300 mm × 448 mm having a gap of 5 mm between the boxes. The needle linear bearing of 5 mm thickness is placed between the shear boxes to maintain the 5 mm gap between the boxes and avoid any frictional forces during shearing to the sample. The upper shear box is fixed and the joint is sheared by moving the lower box. The upper box can only move in the vertical direction. The lower box is fixed on a rigid base through bearings, which can move only in horizontal direction. The apparatus is fitted with two actuators for load and six LVDTs for displacement



Fig. 1 Photograph of large scale direct shear testing machine

measurements. The normal and shear load is applied uniformly on the shearing plane through the load cell of the capacity of 500 and 1000 kN respectively. Displacements are monitored and measured through six number of LVDTs. Four LVDTs are used to measure the normal displacement and to provide a check on specimen rotation about an axis parallel to the shear zone and perpendicular to the shearing direction. Degree of joint closure and dilation angle can also be obtained from these measurements. Two LVDTs are used to measure the shear displacement. These displacement devices have adequate ranges of travel to accommodate the displacements, ±20 mm. Sensitivities of these devices are 0.001 mm for both normal displacement and shear displacement.

The second unit of apparatus consists of hydraulic power pack with servo valve, the function of the hydraulic power pack is to supply required flow and pressure for the actuation of the actuator and the function of servo valve is to control the flow of the hydraulic fluid to the actuator.

Data acquisition and controlling unit is the third unit of apparatus. Data acquisition system consists of a computer installed with direct shear test software and PCI bus advanced data acquisition card. The output of data acquisition system is connected to CPU via cord and the load and deformation values are stored at desired intervals as notepad data. The software is having the capability to control the test, collect the data and plot online graph between shear load versus time, normal load versus time, shear load versus normal load, shear load versus shear displacement, normal load versus normal displacement and shear displacement versus normal displacement. Controlling unit consists of signal conditioning unit and controlling unit. Signal conditioning unit receives the output signal from the various transducers which amplifies and process the signal as per the requirement and transfer it to computer through connecting cables where it is accepted by the data acquisition system. The detailed working of this equipment is presented in Shrivastava and Rao (2013).

In this apparatus, CNL and CNS conditions are achieved by an electro hydraulic servo-valve, which controls the application of hydraulic power to linear actuator to provide the programmed normal force to test specimens. The programmed normal force is calculated from the Eq. (1), as proposed by Shrivastava and Rao (2013). The Eq. (1) is also made the part of direct shear testing software which is used for monitoring, recording and plotting the test results. The asperity angle (i), initial normal load (P_n) and stiffness of the rock joints or reinforcement (k_n) is feeded as input data in the direct shear testing software. The dilation and horizontal displacement data collected through sixteen channel data acquisition system is also feeded to direct shear testing software. The increased load during the progress of testing is calculated by the following equation:

$$P_{n(t+\Delta t)} = P_{n(E)} + k_n(Y - Y') \quad (1)$$

where $P_{n(t+\Delta t)}$ = normal force at any time interval $t + \Delta t$, $P_{n(t)}$ = normal force at any time interval t, k_n = stiffness of the surrounding rock mass, $Y - Y'$ = dilation resisted by the surrounding rock mass, Y = free dilation at any shear deformation.

The typical calculation for a given asperity angle (i) is as shown in Fig. 2, Y = X tani when sample is

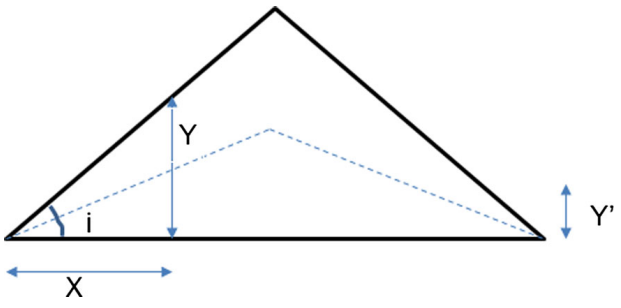


Fig. 2 Typical asperity profile

free to dilate, from the given JRC value the equivalent *i* value can be calculated by the method suggested by Xie and Pariseau (1992) and Maksimovic (1996), X = horizontal displacement measured as the average of two horizontal LVDT's placed near the sample, Y′ = dilation measured as the average of readings of four normal LVDT's placed on the top of the sample.

3 Physical Modeling

3.1 Selection of Model Material

It is difficult to interpret the result of direct shear test on natural rock because of difficulties in repeatability of the sample. To overcome this problem a model material is selected which can be easily handled and reproducibility of the sample can be ensured. Plaster of Paris (POP) is selected because of its universal availability and its modulability into any shape when mixed with water to produce the desired joints and also long term strength is independent of time once the chemical hydration is completed.

To perform a series of physical and mechanical tests a number of specimens were prepared by mixing the prescribed quantity of water with POP powder. The prescribed percentage of water is decided so as to achieve proper workability of the paste and required strength to simulate soft rock. The different water powder ratio is tried and the desired strength and workability is obtained with water powder ratio of 0.60. The resultant paste after 1–2 min of mixing was poured into aluminum cylindrical moulds. The sample is removed from the mould after about 30–40 min of pouring, so that sample gets sufficiently hardened and stand freely for curing in air. Now, the specimens were air cured at room temperature for 14 days before testing. Initial size of the specimen was 38 mm in diameter and 76 mm in height, which is used for unconfined compressive strength tests. The samples for other tests are cut from this samples depending upon size requirements.

The basic properties of the model material like dry density (Υ_d), uniaxial compressive strength (σ_c), Poisson's ratio (ν) and tangent modulus (E_{t50}) were determined in the laboratory as per the suggested methods ISRM (1977, 1979). The basic properties of the model material at 60 % of the moisture are presented in Table 1. The average uniaxial compressive strength of model material is 11.75 MPa and average tangent modulus at 50 % of peak axial stress is estimated 2281 MPa. Thus, the material can be classified as 'EL' based on Deere and Miller (1966) classification chart, indicating that the material is of very low strength (E) and low modulus ratio (L) and is suitable for simulating the behaviour of jointed rocks like siltstone, sandstone, friable limestone, clay shale and mudstone.

3.2 Sample Preparation

The sample with different asperity and roughness is prepared with the help of specially designed and fabricated casting mould and asperity plate, both is of cast iron. The rock joint is seldom plane, it always contains roughness and joint roughness can be of any shape and size. The roughness can vary within the joints or it will be different for different joints. The physical modelling of joint roughness as it appears in the in situ rock joints is very difficult. Hence, in the

Table 1 Properties of model material

Property	Sample description	No. of samples	Representative value	Testing method
Dry density, Υ_d (kN/m^3)	38 mm diameter, 76 mm height cylinder	20	12.34	ISRM (1977)
Uniaxial compressive strength (intact cylinder), σ_c (MPa)	38 mm diameter, 76 mm height cylinder	12	11.75	ISRM (1979)
Poisson's ratio, ν	38 diameter, 76 mm height cylinder	3	0.22	ISRM (1979)
Tangent modulus, E_{t50} (MPa)	38 mm diameter, 76 mm height cylinder	3	2281	ISRM (1979)
Modulus ratio (E_{t50}/σ_{ci})	–	3	194	ISRM (1979)
Failure strain in UCS, ε_f (%)	38 mm diameter, 76 mm height cylinder	3	0.63	ISRM (1979)
Deere–Miller (1966) classification	–	–	EL	

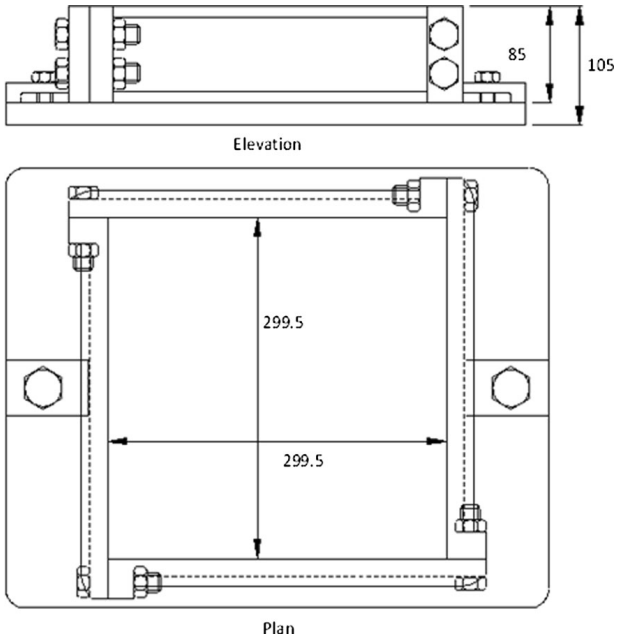


Fig. 3 Schematic diagram of casting mould

present work joint of equal and unequal angle asperity is prepared with the help of specially designed casting mould size 299.5 × 299.5 × 85 mm (Fig. 3) and asperity plate of angles 0°–0°, 30°–30°, 15°–15° and 30°–15°. The schematic diagram and photograph of one of the asperity plate of angle 15°–15° is shown in Fig. 4. These asperity angles can easily be converted into joint roughness coefficient (JRC) value by the use of method suggested by Xie and Pariseau (1992) and Maksimovic (1996).

The two parts namely upper and lower, parts of the joints are prepared simultaneously to create proper mated joints. The casting moulds with suitable platform is placed and screwed properly on the vibrating table. The desired asperity plate is placed inside the mould facing the inclined asperity surface upwards. POP mixed with water thoroughly for 2 min before pouring into the casting mould, by maintaining the water POP ratio as 0.60 and the paste is then vibrated on vibrating table for 1 min to compact and to remove any entrapped air. The top surface of the specimen is properly leveled and excess paste is removed. The samples require 55 min to achieve final setting; hence, sample is removed from the mould only after elapse of 55 min after adding water into the POP, so as to insure proper shape of the sample. The samples are allowed to cure in air for 14 days before testing. The different roughness conditions of the rock joints were simulated by preparing four types of specimen sets namely 0°–0°, 30°–30°, 15°–15° and 30°–15°, the sample with asperity 30°–30°,15°–15° and 30°–15° are presented in Fig. 5.

3.3 Experimental Investigation

The shear behaviour of joints under CNL condition was investigated before joints were tested under CNS

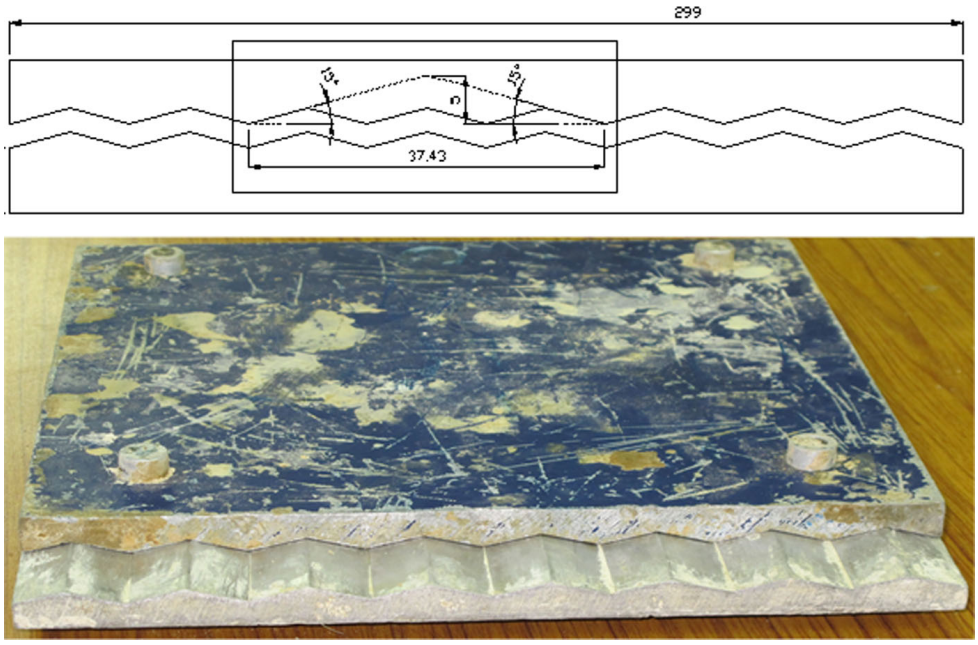


Fig. 4 Schematic diagram and photograph of asperity 15°–15° (all dimensions are in mm)

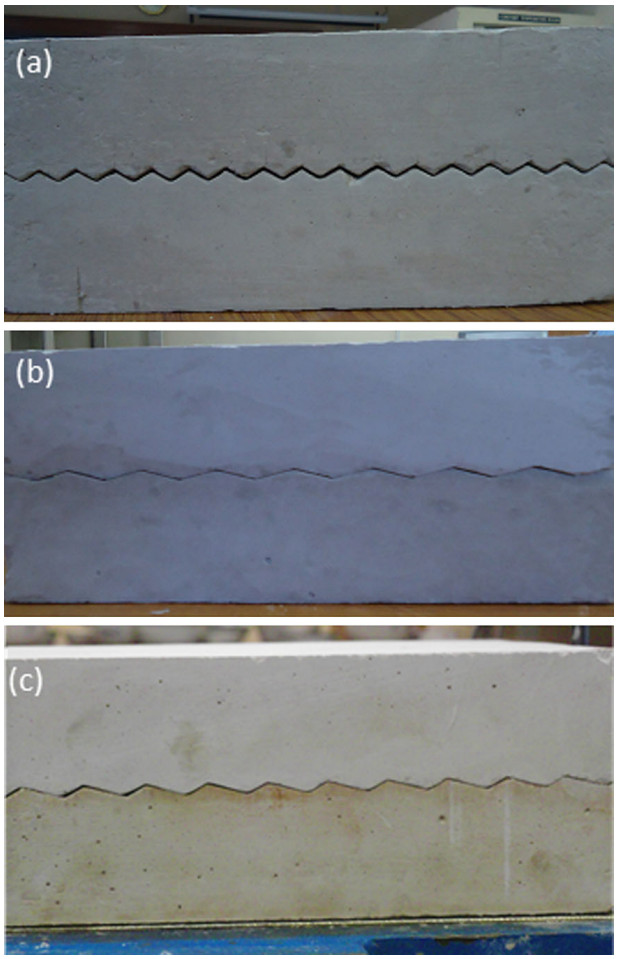


Fig. 5 Photograph of samples with asperity 30°–30°, 15°–15° and 30°–15°

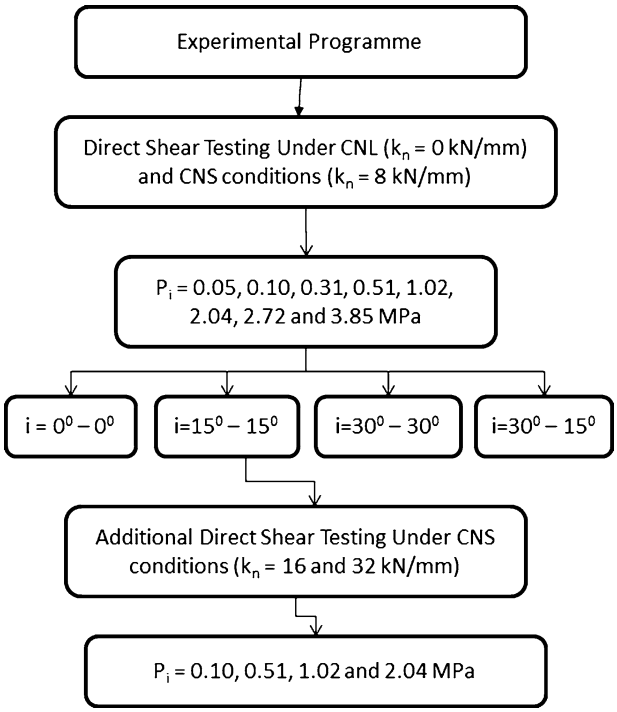


Fig. 6 Experimental programme

conditions. This has been carried out to assess the affect of boundary conditions on the shear behaviour. Hence, a series of tests were performed on the physically modeled rock joints with asperity angles 30°–30°, 15°–15° and 30°–15°. To study the affect of the asperity on shear behaviour, tests were also performed on planar joints with asperity angle 0°–0° for the boundary conditions, complete test, summary of the tests programme is presented as flow chart in Fig. 6.

In the present study 0.50 mm/min rate of shearing is selected, which is based on the study carried out by Shrivastava (2012) on similar type of physically modeled rock joints, where it has been concluded that there is no effect of shearing rate on peak shear stress up to shearing rate of 0.5 mm/min and at shearing rate more than 0.5 mm/min the effect is to increase the peak shear stress for both the boundary conditions. The effect of CNS boundary conditions on shear behaviour is discussed in detail by Shrivastava and Rao (2010, 2011) and Shrivastava et al. (2011).

3.3.1 Shear Behaviour

The shear behaviour of 30°–30°, 15°–15° and 30°–15° asperity joint under CNL ($k_n = 0$ kN/mm) and CNS ($k_n = 8$ kN/mm) boundary condition is plotted as shown in Figs. 7, 8 and 9. To get the normal and shear stress normal load and shear load is divided by the initial cross sectional area of the sample. The stress–displacement behaviour is characterized by a well defined peak. It is clear from the test result that CNL boundary condition always under predicts the shear strength of the joint as compared to CNS boundary condition for the same initial normal stress. The shear stress response for planar joints under various P_i is presented in Fig. 10. The stress–displacement behaviour is characterized by increase in the shear stress with shear displacement till peak is reached and thereafter shear stress is found to be constant. The peak is reached when the shear stress almost equal to the normal stress acting on the joint. The constant shear stress indicates sliding of the sample on the failure surface. The CNS condition is not possible for

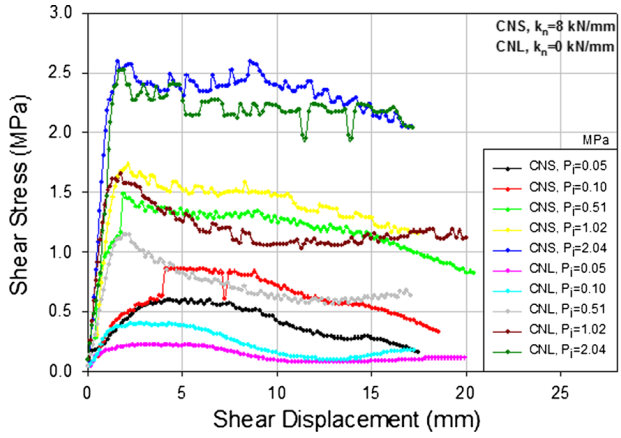


Fig. 7 Shear behaviour of 30°–30° asperity joint under CNL and CNS boundary condition

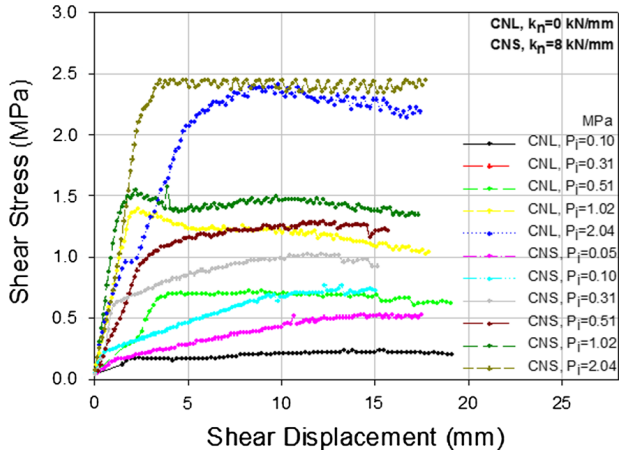


Fig. 8 Shear behaviour of 15°–15° asperity joint under CNL and CNS boundary condition

planar joint because for planar joint Y = 0 and $P_{n(t)}$ will be equal to $P_{n(t+\Delta t)}$, which can be seen by substituting these values in Eq. (1).

Effects of boundary conditions on the shear strength of 0°–0° rock joints are compared with rock joints having regular i.e. 15°–15°, 30°–30° and irregular i.e. 30°–15° triangular asperities. The percentage (%) increase in shear strength for different asperity at different P_i for both CNL and CNS conditions are plotted in Fig. 11. It is observed that the shear strength increases with increase in asperity angle because of increased frictional resistance offered by the asperity. But the % increase in peak shear stress decreases with increase in P_i for both CNL and CNS conditions, it is due to degradation in asperity under that normal stress. The percentage of increase in shear strength from 0° to 0° joint to 15°–15° and 30°–

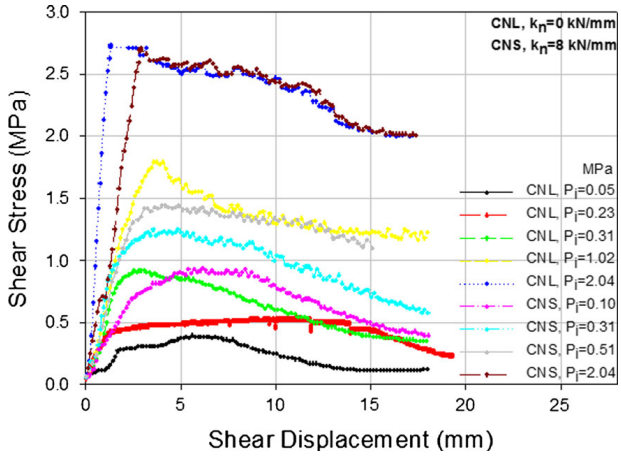


Fig. 9 Shear behaviour of 30°–15° asperity joint under CNL and CNS boundary condition

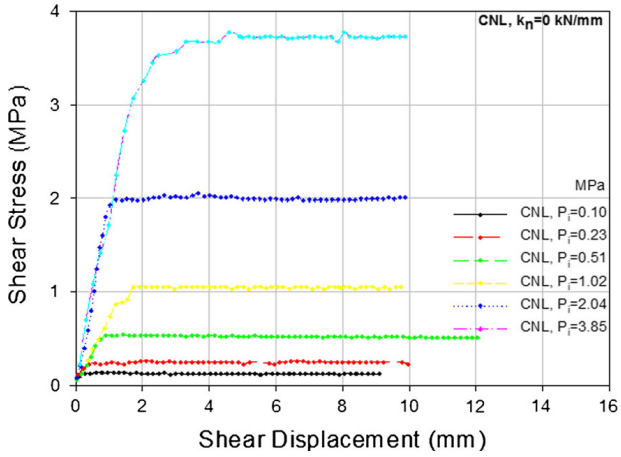


Fig. 10 Shear behaviour of planar joint

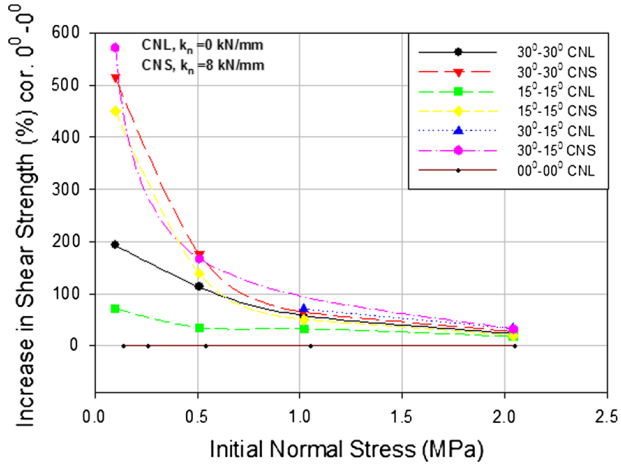


Fig. 11 Increase in shear strength under CNS conditions for different asperity with increasing P_i

30° are 71.42 and 192.85 respectively for CNL conditions at $P_i = 0.10$ MPa. It reduces to 17.56 and 23.41 % when P_i increased to 2.04 MPa for the above condition. Similarly, for CNS condition the percentage increase in shear strength from 0° to 0° to 15° to 15°, 30° to 30° and 30° to 15° are 450.00, 514.28, and 571.43 respectively for $P_i = 0.10$ MPa. The above% increase reduces to 19.51, 23.41 and 32.19 respectively when P_i is increased to 2.04 MPa.

The CNS condition shows a strain softening behaviour where as CNL condition strain hardening behaviour. The strain hardening and softening behaviour for CNL and CNS boundary condition is seen when the sheared sample crosses the peak of the asperity i.e. for 30°–30° asperity joint at shear displacement more than 8.66 mm as shown in Fig. 7. The strain softening behaviour for the CNS condition is due to decrease in normal stress during shearing when the sheared sample crosses the peak of the asperity. The strain hardening behaviour for the CNL conditions can be attributed to deposition of the sheared material within depressed portion of the joints, which has resulted into increase in shear stress.

The normal stress on the shear plane remains constant during testing for CNL conditions. However, for CNS conditions normal stress increases as asperity slides one above the other. The results show that the normal stress increases with shear displacement and reaches to the maximum value at shear displacement near to peak of the asperity i.e. 8.66 mm and after that the normal stress decreases to reach to the minimum value which is equal to initial normal stress at shear displacement equal to length of one asperity i.e. 17.32 mm as shown in Fig. 12 for 30°–30° asperity. In the CNS boundary condition variation of normal stress follows the shape of the asperity. The percentage increase in normal stress is very high about 900 % at low P_i i.e. $P_i = 0.10$ MPa and at $P_i > 0.51$ MPa the increase in normal stress is negligible for the entire asperity angle as depicted in Fig. 13.

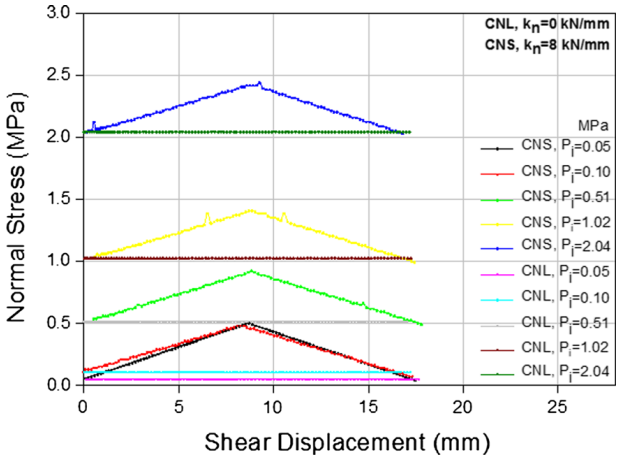


Fig. 12 Variation of normal stress of 30°–30° joints under CNL and CNS condition

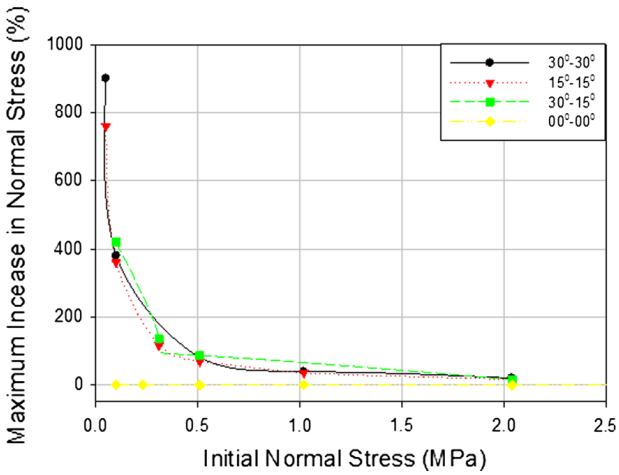


Fig. 13 Effect of asperity in % increase normal stress under CNS conditions

4 New Shear Strength Model

4.1 Back Ground for Development of New Model

Classical theory proposed by Newland and Allely (1957) and Withers (1964) indicates that roughness along with the joint surface plays a vital role in contributing the shear strength of rock joints. Based on a series of direct shear tests performed under CNL conditions on artificial joint specimens with regular teeth inclinations at different P_i, Patton (1966) proposed the following bilinear strength model.

$$\tau = P_i \tan(\varphi_b + i) \tag{2}$$

where P_i, Φ_b and i are the initial normal stress, basic friction angle and asperity angle of the joint surface respectively.

The above Eq. (2) is valid when dilation is not restricted, joint is subjected to low normal stress and degradation of the joint surface does not take place during shearing.

If dilation is inhibited and normal stress is high the degradation of asperities occurs and shearing will take place across the asperity. This means asperity angle

changes with normal stress for CNL condition. Therefore, some researchers in the past tried to relate reduction in asperity angle with normal stress and other joint parameters. Instead of dilation angle in Eq. (2), Barton (1973) and Maksimovic (1996) has proposed the term JRC log (JCS/σ_n) and $\Delta\Phi(1 + \sigma_n/P_n)$ respectively. But these modifications do not account for change in normal stress during the shearing process for CNS conditions.

4.2 Development of New Shear Strength Model

The shear stress and shear displacement behaviour of modelled rock joint can be divided into three zones, different zones of one typical 30°–30° asperities is presented in Fig. 14. In zone I predominantly sliding of the sample take place without shearing of the asperity. The limit of the zone-I depends upon the shear strength of the material and shear stress increases at higher rate with small shear displacement in this zone. The angle of friction of the model material is 45°, which limits the zone-I at shear stress = P_i. In zone-II, shearing of the asperity is more predominant than the sliding. The limit of the zone-II is up to maximum shear stress, in this zone rate of increase in shear stress decreases with shear displacement. Zone-III is the last zone where all the asperity is sheared off. Due to deposition of the crushed material on the joints, shear stress decreases or increases slightly with shear displacement depending upon CNL or CNS conditions.

The shear test results on planar rock joint i.e. 0°–0° asperity modelled rock joint reflects that the strength envelope for the CNL boundary condition is linear for all range of initial normal stress. Where as for non planar rock joint i.e. 15°–15°, 30°–15° and 30°–30° asperities the strength envelope is curvilinear for both CNL and CNS boundary condition and the curvature of the strength envelope increases with increase in asperity angle as presented in Fig. 15. The curvature of the strength envelopes also change with change in the initial normal stress for non planar joint. The curvature of the strength envelope is same up to low normal stress i.e. $P_i \leq 0.09\ \sigma_c$ and after that the curvature of the strength envelope is increased and approaches more towards the linearity for both CNL and CNS conditions. At $P_i > 0.09\ \sigma_c$ and $P_i < 0.18\ \sigma_c$, the shear strength increases with increase in the asperity angle. But increase in the P_i reduces the effect of the asperity angle on increasing the shear strength. The reasons for above are degradation of the asperity angle at high normal stress and joints behave almost like a planar joint. At high initial normal stress i.e. $P_i > 0.18\ (\sigma_c)$, there is no effect of the asperity angle and boundary condition on the shear strength.

On the basis of experimental observations and results, a shear strength model is developed. Bilinear shear strength model proposed by Patton (1966) is used as a basic equation, for the development of the present model. The two major limitations of Eq. (2) is presented below:

1. unable to use correct value of P_i for CNS conditions,
2. the effect of asperity degradation is not considered due to increase in normal stress.

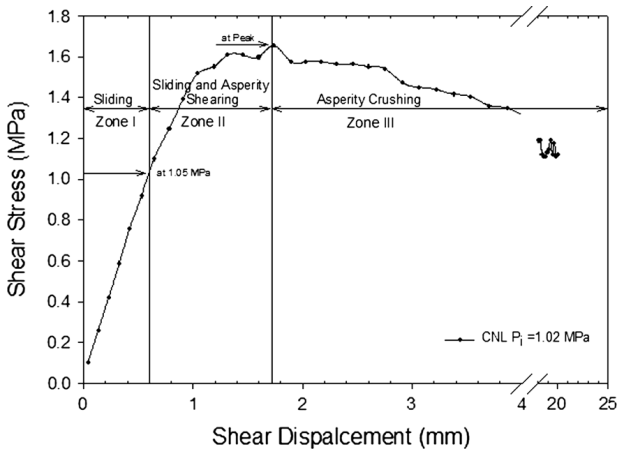


Fig. 14 Different zones of shearing

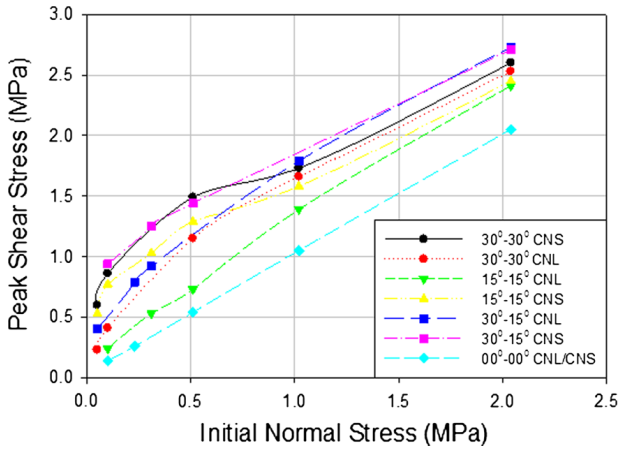


Fig. 15 Comparison of strength envelopes of different asperity (CNL and CNS)

Hence, Eq. (2) is modified to overcome the above limitations. It is developed by assuming that at peak shear stress under CNS condition, normal stress momentarily remains constant. The equation is as given below:

$$\tau_p = P_n \tan(\varphi_b + i') \tag{3}$$

where τ_p = peak shear stress in MPa, P_n = normal stress corresponding to peak shear stress in MPa for CNL/CNS condition, φ_b = basic friction angle (°), i' = effective asperity angle (°).

4.2.1 Prediction Model for Normal Stress (P_n)

The increase in normal stress under CNS conditions is governed by the stiffness of the surrounding rock joint and dilation resisted during shearing. The linear variation of the normal stress is observed during shearing in the present study and similar observation is made by Jiang et al. (2004), Shrivastava and Rao (2010, 2011), Shrivastava et al. (2011) and Shrivastava (2012). The variation of normal stress corresponding to peak shear stress with P_i for different stiffness and asperity angles are presented in Fig. 16. The data fitting of the graph indicates that there is a linear relationship between the normal stress corresponding to peak shear stress and P_i as shown by Eq. (4).

$$P_n = aP_i + b \tag{4}$$

where P_n is normal stress corresponding to peak shear stress, P_i is initial normal stress in MPa, a and b are constant which depend upon the asperity angle (i) and normal stiffness (k_n).

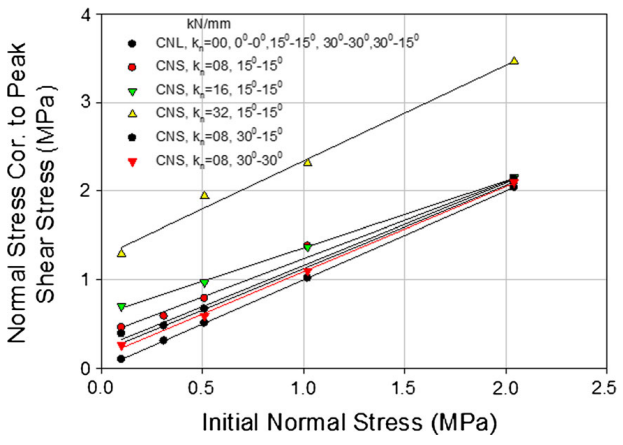


Fig. 16 Variation of normal stress for different asperity angle

The coefficient a and b along with coefficient of determination (R^2) is given in Table 2, it can be seen that constant 'a' is almost insensitive to k_n and asperity angle and hence a = 1 is used for all the conditions. But, the coefficient b is sensitive to both k_n and i, it decreases with increase in asperity angle and linear relationship exist between coefficient, b and k_n. The increase in asperity angle causes increase in dilation and reduction in dilation resited which in turn reduces the increase in normal stress. Hence the generalized coefficients are:

$$a = 1 \tag{5}$$

$$b = 0.6(k_n/i) \tag{6}$$

where k_n is normal stiffness of the joint in kN/mm, i is the asperity angle (°).

The experimentally determined normal stress (P_n) is compared with the proposed model of different asperity joint at different normal stiffness conditions and the variations of results are presented in Fig. 17. It can be seen that variation of predicted results are within 95 % of prediction band.

4.2.2 Prediction Model for Effective Asperity Angle (i')

The fine observation of the sheared samples shows that at low normal stress, sliding of the sample takes place with shear displacement and increase in normal stress causes degradation of the asperity as discussed in Sects. 3.3.1 and 4.2. The rate of asperity degradation depends upon the ratio of P_n/σ_c. Increase in this ratio causes flattening of the asperity and the reduction in the effective asperity angle. The decay rate of asperity angle is exponential as shown in Fig. 18, the figure shows that most of the experimental data falls within the 95 % confidence band. This can be represented by the following equation:

$$\frac{i'}{i} = 1.25e^{-9.86\frac{P_n}{\sigma_c}} \tag{7}$$

where i' = effective asperity angle, i = initial asperity angle, P_n = normal stress corresponding to peak shear stress, σ_c = uniaxial compressive strength in MPa.

Now, P_n and i' are calculated from the Eqs. (4) and (7) respectively. These values are substituted in the Eq. (3) to predict the shear strength under CNL and

Table 2 Coefficients for calculation of normal stress

Asperity angle (i) (°)	Normal stiffness k_n (kN/mm)	a	b	R^2
00–00	0 (CNL)	1	0	1
15–15	0 (CNL)	1	0	1
	8 (CNS)	0.865	0.373	0.988
	16 (CNS)	0.753	0.606	0.999
	32 (CNS)	1.081	1.258	0.989
30–30	0 (CNL)	1	0	1
	8 (CNS)	0.958	0.133	0.998
30–15	0 (CNL)	1	0	1
	8 (CNS)	0.929	0.232	0.996

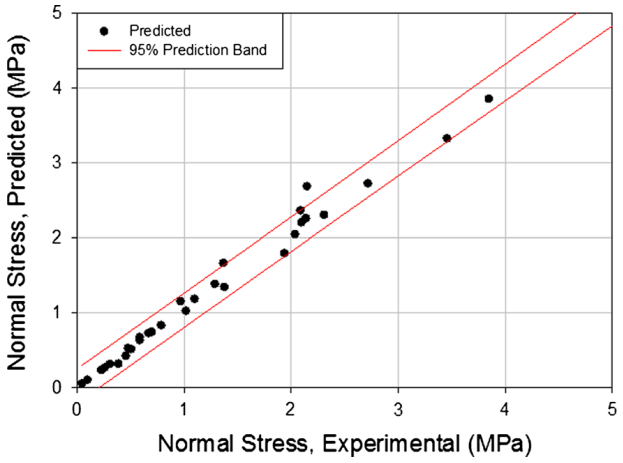


Fig. 17 Variation of predicted normal stress (P_n) from experimental results

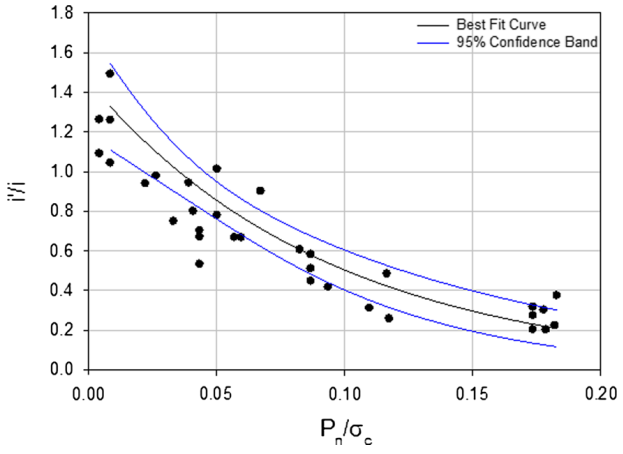


Fig. 18 Asperity decay with increase in normal stress for CNL and CNS conditions

CNS conditions. To predict the basic parameters like Φ_b, σ_c and k_n are required for any rock joints, which can be easily determined by simple testing facility or

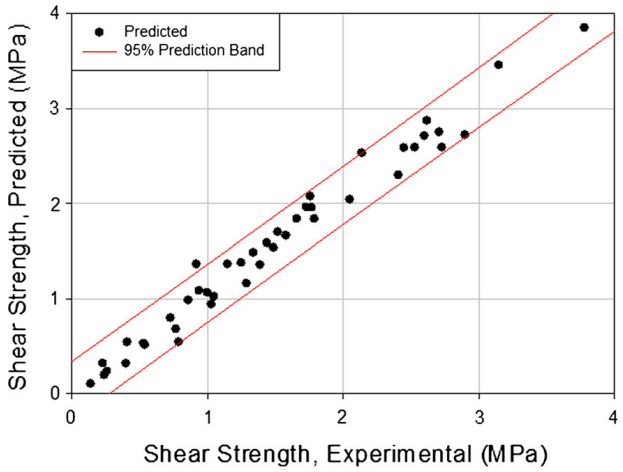


Fig. 19 Variation of predicted shear strength from experimental results

data available in the literature for different types of rock. The asperity angle (i) is either measured or can be calculated from JRC values by the method suggested by Maksimovic (1996), which can be approximated to $i = 2 \times \text{JRC}$. The predicted peak shear stress of proposed model is compared with the experimental results and the variations of the results are plotted in Fig. 19. It can be observed that most of the predictions are within the prediction band of 95 %. The predicted peak stress and normal stress for different boundary conditions is compared with the experimental results and model proposed by Patton (1966) and Barton (1973), which is presented in Table 3. The shear strength predicted by proposed model is closer to experimental value for both CNL and CNS conditions. But shear strength predicted by Patton (1966) and Barton (1973) gives comparable results only in case of CNL conditions (i.e. $k_n = 0$) but

Table 3 Experimental and predicted results of peak shear stress and normal stress

Asperity angle (i) (°)	k_n (kN/mm)	P_i (MPa)	τ_p exp. (MPa)	P_n exp. (MPa)	P_n pred. (MPa)	τ_p pred. (MPa)	τ_p Barton (1973) (MPa)	τ_p Patton (1966) (MPa)
15–15	0	0.10	0.24	0.10	0.10	0.19	0.18	0.17
15–15	0	0.31	0.53	0.31	0.31	0.53	0.47	0.54
15–15	0	0.51	0.73	0.51	0.51	0.79	0.73	0.88
15–15	0	1.02	1.39	1.02	1.02	1.35	1.35	1.77
15–15	0	2.04	2.41	2.04	2.04	2.30	2.49	3.53
15–15	8	0.1	0.77	0.46	0.42	0.68	0.18	0.17
15–15	8	0.31	1.03	0.59	0.63	0.94	0.47	0.54
15–15	8	0.51	1.29	0.79	0.83	1.16	0.73	0.88
15–15	8	1.02	1.58	1.38	1.34	1.66	1.35	1.77
15–15	8	2.04	2.45	2.09	2.36	2.58	2.49	3.53
15–15	16	0.1	1.00	0.70	0.74	1.06	0.18	0.17
15–15	16	0.51	1.34	0.97	1.15	1.48	0.73	0.88
15–15	16	1.02	1.77	1.37	1.66	1.95	1.35	1.77
15–15	16	2.04	2.62	2.15	2.68	2.87	2.49	3.53
15–15	32	0.1	1.52	1.29	1.38	1.70	0.18	0.17
15–15	32	0.51	1.76	1.94	1.79	2.07	0.73	0.88
15–15	32	1.02	2.14	2.31	2.30	2.53	1.35	1.77
15–15	32	2.04	3.15	3.46	3.32	3.46	2.49	3.53
30–30	0	0.05	0.23	0.05	0.05	0.31	0.30	0.19
30–30	0	0.1	0.41	0.1	0.1	0.54	0.40	0.37
30–30	0	0.51	1.15	0.51	0.51	1.36	1.12	1.90
30–30	0	1.02	1.66	1.02	1.02	1.84	1.83	3.81
30–30	0	2.04	2.53	2.04	2.04	2.59	3.07	7.61
30–30	8	0.1	0.86	0.26	0.26	0.98	0.40	0.37
30–30	8	0.51	1.49	0.59	0.67	1.53	1.12	1.90
30–30	8	1.02	1.73	1.1	1.18	1.96	1.83	3.81
30–30	8	2.04	2.60	2.1	2.2	2.71	3.07	7.61
30–15	0	0.05	0.40	0.05	0.05	0.31	0.30	0.19
30–15	0	0.23	0.79	0.23	0.23	0.92	0.65	0.86
30–15	0	0.31	0.92	0.31	0.31	1.07	0.79	1.16
30–15	0	1.02	1.79	1.02	1.02	1.84	1.83	3.81
30–15	0	2.04	2.73	2.04	2.04	2.59	3.07	7.61
30–15	8	0.1	0.94	0.39	0.26	0.98	0.40	0.37
30–15	8	0.31	1.25	0.48	0.47	1.31	0.79	1.16
30–15	8	0.51	1.73	0.67	0.67	1.53	1.12	1.90
30–15	8	2.04	2.71	2.14	2.2	2.71	3.07	7.61
0–0	0	0.1	0.14	0.1	0.1	0.10	0.10	0.10
0–0	0	0.23	0.26	0.23	0.23	0.23	0.23	0.23
0–0	0	0.51	0.54	0.51	0.51	0.51	0.51	0.51
0–0	0	1.02	1.05	1.02	1.02	1.02	1.02	1.02
0–0	0	2.04	2.05	2.04	2.04	2.04	2.04	2.04
0–0	0	2.72	2.9	2.72	2.72	2.72	2.72	2.72
0–0	0	3.85	3.78	3.85	3.85	3.85	3.85	3.85

Table 4 Validation of proposed model for different rock samples (CNS)

Description of sample	σ_c (MPa)	k_n (kN/mm)	i (°)	P_i (MPa)	τ_p exp. (MPa)	τ_p prop. (MPa)	References
Gypsum plaster	20.0	15.0	4.6	0.35	0.72	0.83	Jeong et al. (2010)
	20.0	15.0	4.6	0.70	0.95	1.11	
	20.0	15.0	4.6	1.05	1.35	1.39	
Resin concrete	107.7	108.0	4.5	2.00	4.22	3.66	Jiang et al. (2004)
	107.7	108.0	4.5	5.00	5.59	6.10	
Johnstone	3.5	26.9	10.0	0.40	0.86	0.68	Seidel and Haberfield (2002)
	3.5	35.9	5.0	0.40	0.48	0.52	
	3.5	35.9	10.0	0.40	0.64	0.52	
	3.5	35.9	15.0	0.40	0.72	0.53	
Plaster of Paris	12.0	8.5	18.5	0.05	1.30	0.88	Indraratna et al. (1998)
	12.0	8.5	18.5	0.16	1.36	1.05	
	12.0	8.5	18.5	0.30	1.86	1.34	
	12.0	8.5	18.5	0.56	2.25	1.85	
	12.0	8.5	18.5	1.10	2.44	2.31	
	12.0	8.5	18.5	1.64	3.12	3.01	
	12.0	8.5	18.5	2.43	0.49	0.66	
	12.0	8.5	9.5	0.16	0.66	0.82	
	12.0	8.5	9.5	0.30	1.01	1.10	
	12.0	8.5	9.5	0.56	1.54	1.64	
	12.0	8.5	9.5	1.10	1.80	2.14	
	12.0	8.5	9.5	1.63	2.72	2.89	
	12.0	8.5	9.5	2.43	1.07	1.02	
	12.0	8.5	26.5	0.05	1.14	1.18	
	12.0	8.5	26.5	0.16	1.61	1.35	
	12.0	8.5	26.5	0.30	1.68	1.61	
	12.0	8.5	26.5	0.56	2.05	2.07	
	12.0	8.5	26.5	1.10	2.82	2.48	
	12.0	8.5	26.5	1.63	3.35	3.12	
	12.0	8.5	26.5	2.43	0.86	0.68	

for CNS conditions (i.e. $k_n > 0$) the results are under predicted.

The Proposed model is validated with some of the experimental results available on natural and artificial rock joints in the literature. As the proposed model require basic parameters like Φ_b, σ_c, and i, where ever these parameters are not reported in the experimental results by different researcher than, for JRC value equivalent asperity angle (i) is calculated by the method suggested by Maksimovic (1996) and for Φ_b, σ_c the value available in the literature for different types of rock is used. Shear strength reported in literature based on experimentation for different types of rock or model material is also predicted by proposed Eq. (3), comparison of results are presented in Tables 4 and 5 for CNS and CNL conditions respectively, the predicted results are in very close agreements with experimental value for both the boundary conditions. The regression analysis has been done for the results presented in Table 4 and presented in Fig. 20, it can be observed that the proposed model has predicted the results within the prediction band of 95 %.

Table 5 Validation of proposed model for different rock samples (CNL)

Description of sample	Φ_b (°)	σ_c (MPa)	P_i (MPa)	JRC	τ_P prop. (MPa)	τ_p exp. (MPa)	References
Limestone 1 (rough fracture)	35.7	26	0.50	17.1	1.46	1.08	Asadollahi and Tonon (2010)
	35.7	26	1.50	14.3	2.22	2.53	
Sandstone (rough fracture)	31.4	41	0.50	15.0	1.06	0.82	
	31.4	41	1.00	8.9	1.15	1.00	
	31.4	41	1.50	5.9	1.34	1.23	
Granite (rough fracture)	24.3	138	1.00	16.2	1.88	1.56	
	24.3	138	1.50	17.7	3.08	2.33	
	24.3	138	2.00	19.4	4.57	3.19	
Limestone 2 (rough fracture)	39.2	155	0.50	9.1	0.91	0.92	
	39.2	155	1.00	9.3	1.81	1.67	
	39.2	155	2.00	10.1	3.67	3.17	
Limestone 1 sawed	35.7	26	1.00	1.0	0.76	0.75	
	35.7	26	5.00	1.0	3.64	3.52	
	35.7	26	0.50	1.0	0.39	0.33	
	35.7	26	1.00	1.0	0.76	0.73	
	35.7	26	2.00	1.0	1.50	1.54	
	35.7	26	4.00	1.0	2.93	3.10	
	35.7	26	5.00	1.0	3.64	3.50	
	35.7	26	3.00	1.0	2.22	2.10	
Sandstone sawed	31.4	41	0.50	1.0	0.33	0.39	
	31.4	41	1.00	1.0	0.66	0.76	
	31.4	41	1.50	1.0	0.98	0.88	
	31.4	41	0.60	1.0	0.40	0.42	
	31.4	41	1.20	1.0	0.79	0.79	
	31.4	41	2.00	1.0	1.30	1.40	
Granite sawed	24.9	138	4.00	1.0	2.02	1.79	
	24.9	138	2.50	1.0	1.27	1.35	
Limestone 2 sawed	39.2	155	0.60	1.0	0.53	0.58	
	39.2	155	0.80	1.0	0.71	0.71	
	39.2	155	0.70	1.0	0.62	0.59	
	39.2	155	0.90	1.0	0.80	0.72	
	39.2	155	1.10	1.0	0.97	0.83	
	39.2	155	1.30	1.0	1.15	0.98	
Gneiss	36.0	184	1.90	12.9	4.10	3.40	Grasselli and Egger (2003)
	36.0	184	3.52	7.4	4.40	4.00	
	36.0	184	3.57	6.7	4.23	3.90	
	36.0	184	3.52	8.6	4.81	4.30	
	36.0	184	4.08	1.8	3.38	3.30	
	36.0	184	2.60	9.4	3.92	3.50	
	36.0	87	0.87	13.0	1.90	1.71	
Granite	34.0	173	1.12	14.2	2.68	2.40	
	34.0	173	1.12	15.9	3.31	2.90	
Gneiss	34.0	173	2.65	3.3	2.31	2.40	

Table 5 continued

Description of sample	Φ_b (°)	σ_c (MPa)	P_i (MPa)	JRC	τ_P prop. (MPa)	τ_p exp. (MPa)	References
Concrete replica	30.0	30	0.035	0.0	0.02	0.02	Desai and Fishman (1991)
	30.0	30	0.138	0.0	0.08	0.08	
	30.0	30	0.345	0.0	0.20	0.20	
	30.0	30	0.035	0.0	0.02	0.02	
	30.0	30	0.069	0.0	0.04	0.04	
	30.0	30	0.138	0.0	0.08	0.09	

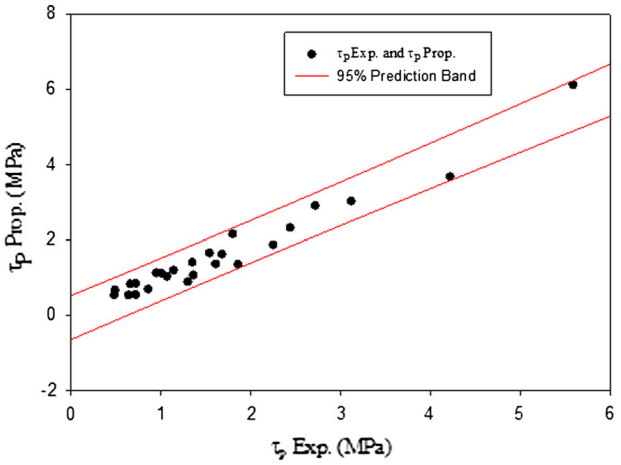


Fig. 20 Variation of predicted shear strength from results available in the literature for CNS conditions

5 Conclusions

Direct shear tests have been performed on physically modeled rock joints with asperity angle 30°–30°, 15°–15°, 30°–15° and 0°–0° under CNL and CNS conditions at different P_i to study the effect of normal stiffness and roughness of the rock joints on shear behavior. It is observed that normal load on the shearing plane is constant under CNL conditions and it increases under CNS conditions during shearing process. The variation of normal stress follows the profile of the asperity angle under CNS condition. The effect of asperity angle is to increase the shear strength with increase in the asperity angle and it increases more for CNS than CNL conditions, but the effect is reducing with increase in P_i. At $P_i \geq 0.18\ \sigma_c$ the shear strength under CNL and CNS conditions are almost same. Asperities with irregular profile have higher shear strength than the regular profile. The strength envelope for planar joint is linear and it changes to curvilinear for non planar joints under both CNL and CNS condition. Based on the series of tests conducted on different asperities joints under CNL and CNS conditions a new model for predicting the shear strength of rock joints is proposed. The proposed model is validated by predicting the shear strength of all types of rocks ranging from soft to hard under CNL and CNS conditions for different asperity angle or JRC value, whose experimental results are available in the literature and it is found that the proposed model successfully describes the shear strength of all the joints under both CNL and CNS conditions.

References

Asadollahi P, Tonon F (2010) Constitutive model for rock fractures: revisiting Barton's empirical model. Eng Geol 113:11–32

Barton N (1973) Review of a new shear strength criterion for rock joints. Eng Geol 7:287–332

Byerlee JD (1975) The fracture strength and frictional strength of weber sandstone. Int J Rock Mech Min Sci Geomech Abst 12:1–4

Deere DU, Miller RP (1966) Engineering classification and index properties of rock. Technical report no. AFNL-TR-65-116. Air Force Weapons Laboratory, Albuquerque

Desai CS, Fishman KL (1991) Plasticity-based constitutive model with associated testing for joints. Int J Rock Mech Min Sci Geomech Abst 28:15–26

Ghazvinian AH, Taghichian A, Hashemi M, Mar'ashi SA (2010) The shear behaviour of bedding planes of weakness between two different rock types with high strength difference. Rock Mech Rock Eng 43:69–87

Goldestin M, Gooser B, Pyrogovsky N, Tulinov R, Turovskaya A (1966) Investigation of mechanical properties of cracked rock, vol 1. In: Proceedings of the 1st congress international society in rock mechanics, Lisbon, pp 521–524

Grasselli G, Egger P (2003) Constitutive law for the shear strength of rock joints based on three-dimensional surface parameters. Int J Rock Mech Min Sci Geomech 40:25–40

Indraratna B, Haque A, Aziz N (1998) Laboratory modelling of shear behaviour of soft joints under constant normal stiffness condition. J Geotech Geol Eng 16:17–44

ISRM (1977) Suggested method for determining water content, porosity, density, absorption and related properties and swelling and slake-durability index properties. Pergamon Press

ISRM (1979) Suggested method for determining the uniaxial compressive strength and deformability of rock materials. Pergamon Press

Jeong S, Ahn S, Seol H (2010) Shear load transfer characteristics of drilled shafts socketed in rocks. Rock Mech Rock Eng 43:41–54

Jiang Y, Xiao J, Tanabashi Y, Mizokami T (2004) Development of an automated servo-controlled direct shear apparatus applying a constant normal stiffness condition. Int J Rock Mech Min Sci 41:275–286

Kim DY, Chun BS, Yang JS (2006) Development of a direct shear apparatus with rock joints and its verification tests. Geotech Test J 29(5):1–9

Ladanyi B, Archambault G (1970) Simulation of shear behaviour of a jointed rock mass. In: Proceedings of 11th symposium on rock mechanics: theory and practice. American Institute of Mining, Metallurgical and Petroleum Engineers, New York, pp 105–125

Maksimovic M (1996) The shear strength components of a rough rock joint. Int J Rock Mech Min Sci Geomech Abst 33:769–783

Newland PL, Allely BH (1957) Volume changes in drained triaxial tests on granular materials. Geotechnique 7:17–34

Obert L, Brady BT, Schmechel FW (1976) The effect of normal stiffness on the shear resistance of rock. Rock Mech 8(2):57–72

Ooi LH, Carter PJ (1987) A constant normal stiffness direct shear device for static and cyclic loading. Geotech Test J 10(1):3–12

Patton FD (1966) Multiple modes of shear failure in rock and related materials. Dissertation, University of Illinois, Urbana

Rao KS, Shrivastava AK, Singh J (2009) Development of an automated large scale direct shear testing machine for rock. In: Indian geotechnical conference, pp 238–244

Saiang D, Malmgren L, Nordlund E (2005) Laboratory tests on shortcrete rock joints in direct shear, tension and compression. Rock Mech Rock Eng 38:275–297

Seidel JP, Haberfield CM (2002) A theoretical model for rock joints subjected to constant normal stiffness direct shear. Int J Rock Mech Min Sci Geomech Abst 39:539–553

Shrivastava AK (2012) Physical and numerical modelling of shear behaviour of jointed rocks under CNL and CNS boundary conditions. Dissertation, IIT Delhi

Shrivastava AK, Rao KS (2010) Effect of boundary conditions on shear behaviour of rock joint. In: ISRM international symposium and 6th Asian rock mechanics symposium—advances in rock engineering, New Delhi, India

Shrivastava AK, Rao KS (2011) Shear behaviour of non planar rock joint. In: 14th Asian regional conference on soil mechanics and geotechnical engineering, Hong Kong

Shrivastava AK, Rao KS (2013) Development of a large-scale direct shear testing machine for unfilled and infilled rock joints under constant normal stiffness conditions. Geotech Test J 36(5):670–679

Shrivastava AK, Rao KS, Rathod GW (2011) Shear behaviour of rock under different normal stiffness. In: 12th international congress on rock mechanics (ISRM), Beijing, pp 831–835

Withers JH (1964) Sliding resistance along discontinuities in rock mass. Dissertation, University of Illinois

Xie H, Pariseau WG (1992) Fractal estimate of joint roughness coefficients. In: Proceedings of the international conference on fractal and jointed rock masses, Lake Tahoe, CA, pp 125–131

Yang ZY, Chiang DY (2000) An experimental study on the progressive shear behaviour of rock joints with tooth-shaped asperities. Int J Rock Mech Min Sci 37:1247–1259

Information Technology & Software Engineering

Kaswan, et al., J Inform Tech Softw Eng 2015, 5:1
http://dx.doi.org/10.4172/2165-7866.1000144

Research Article Open Access

Software Reliability Modeling using Soft Computing Techniques: Critical Review

Kaswan KS[1]*, Choudhary S[2] and Sharma K[3]

[1]Research Scholar, Department of Computer Science, Banasthali University, Banasthali, Rajasthan, India
[2]Department of Computer Science, Banasthali University, Rajasthan, India
[3]Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

Abstract

To obtained solutions to problems quickly, accurately and acceptably, a large number of soft computing techniques have been developed, but it is very difficult to find out which one is the most suitable and can be used globally. In this paper, we have provided an overview of existing soft computing techniques, and then critically analyzed the work done by the various researchers in the field of software reliability. The probability of failure-free operation of a software system for a specified time in a specifies environment. Further to this, we have also compared soft computing techniques in terms of software reliability modeling capabilities.

Keywords: Neural network; Fuzzy logic; Genetic programming; Cuckoo search; Soft computing and software reliability

Introduction

Software engineering is a discipline whose aim is the production of quality software, software that is delivered on time, within budget, and that satisfies its requirements [1]. Software Engineering are playing very important role in software life and there is always a need of high quality software. Software reliability is the most measurable aspect of software quality. Software reliability can be defined as the probability of failure-free software operation for a specified period of time in a specified environment [2-4]. The software failures are introduced by the system analysts, designers, programmers and managers during different phases of software development life cycle. The probability of failure-free operation of a software system for a specified time in a specified environment. To detect and remove these errors, the software system is tested. The quality of software system in terms of reliability is measured by the removal of these errors. Software reliability modeling plays a significant role in many critical and daily life applications, which has led to the tremendous work being carried out in software reliability modeling. These models successfully have been used for estimation and prediction of the number of errors remaining in the software. User can access the current and future reliability through testing using these models, and can make decisions about the software such as whether the product can be released in its present state or we require further testing in order to improve the quality of software. Soft computing techniques are the collection of different concepts and techniques which aim to overcome the difficulties encountered in real world problems. It deals with the problems that seem to be imprecise, uncertain and difficult to categorize. One may see soft computing as an attempt to mimic natural creatures: plants, animals, human beings, which are soft, flexible, adaptive and clever. In this sense soft computing is the name of a family of problem-solving methods that have analogy with biological reasoning and problem solving. This paper is organized as follows: section II includes different soft computing techniques. In section III we discuss the role of different soft computing techniques in software reliability models. Section IV contains description of soft computing technique in software reliability models. Comparison of these soft computing techniques in terms of modeling capabilities is given in section V. Finally, the paper concludes with section VI.

Soft-computing Techniques

Soft computing is an association of computing methodologies that includes as its principal members' fuzzy logic, chaos theory, neuro-computing, evolutionary computing and probabilistic computing [5]. Soft computing is based on natural as well as artificial ideas. It is referred as a computational intelligence [6]. There are a number of soft computing techniques exists and plays an important role in many areas such as, in computer science, machine learning, artificial intelligence applied in engineering areas such as mobile robot, cooling heating, communication network, inverters, converters, electric power system, power electronics, motion control and aircraft etc.. In this section we have discussed about the classification of existing soft computing techniques as shown in Figure 1.

In Figure 1 we discussed about some soft computing techniques as neural networks, Fuzzy Logic, Support vector machine (SVM), Evolutionary computing, Bayesian Network and Chaos Theory. Then some technique is also used in the combination with the others as Neuro-Fuzzy, the combination of Neural Network and Fuzzy Logic. Evolutionary Computing System is further divided into Evolutionary Algorithm and Swarm Intelligence techniques.

Neural Networks

According to Nigrin A neural network is a circuit composed of a very large number of simple processing elements that are neurally based. Each element operates asynchronously, on local information; thus there is no overall system clock. Applications of neural networks are character recognition, image compression, stock market prediction, traveling salesman's problem, medicine, electronic nose, security and loan applications.

Support vector machine

Boser, Guyon, and Vapnik developed Support Vector Machine (SVM) was introduced, in COLT-92. Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. Support vector machines (SVM) have both a solid mathematical background and good performance in practical

***Corresponding author:** Kaswan KS, Research Scholar, Department of Computer Science, Banasthali University, Banasthali, Rajasthan, India, Tel: 228341; E-mail: KASWANKULDEEP@GMAIL.COM

Received January 29, 2015; **Accepted** March 31, 2015; **Published** April 13, 2015

Citation: Kaswan KS, Choudhary S, Sharma K (2015) Software Reliability Modeling using Soft Computing Techniques: Critical Review. J Inform Tech Softw Eng 5: 144. doi: 10.4172/2165-7866.1000144

Copyright: © 2015 Kaswan KS, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

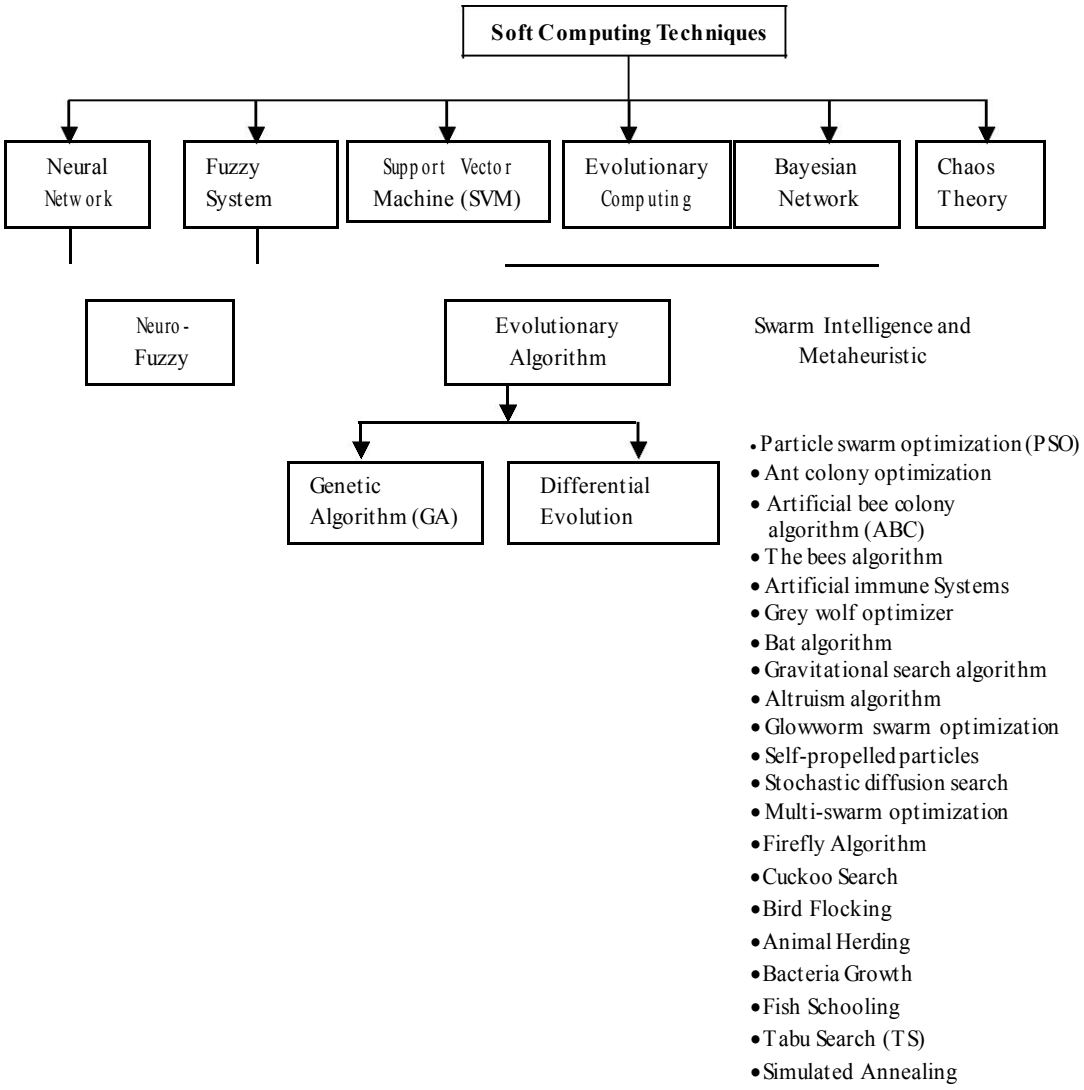


Figure 1: Soft Computing Techniques.

applications, such as image processing, artificial intelligence, medical, pattern recognition, machine learning, applied statistics, business intelligence, and information technology.

Fuzzy logic

Developed by Lotfi A. Zadeh at the University of California in Berkeley. It is a multi-valued logic that allows intermediate values to be defined between conventional evaluations like true/false, yes/no and low/high, etc. [7,8]. The most significant application area of fuzzy logic has been in control field. Fuzzy control having been successfully applied to numerous problems, these includes fans control, complex aircraft engines and control surfaces, wheel slip control, helicopter control, automatic transmission, industrial and missile guidance.

Evolutionary computing

Evolutionary computing can be viewed as an adaptation of a probabilistic approach based on principles of natural evolution [9]. It can also defined as the stochastic search and optimization heuristic approach derived from the classic evolution theory, which are implemented on computers in the majority of cases [10]. Evolutionary algorithms have been successfully applied to numerous problems from different domains, bioinformatics, including optimization, automatic programming, signal processing, social systems [11].

Bayesian network

Bayesian networks are graphical models for reasoning under uncertainty, where the nodes represent variables (discrete or continuous) and arcs represent direct connections between them. various applications such as the impact of management style on statistical efficiency, studies of web site usability, operational risks, biotechnology, customer satisfaction surveys, healthcare systems and the testing of web services.

Chaos theory

A deterministic system is said to be chaotic whenever its evolution sensitively depends on the initial conditions. This property implies that two trajectories emerging from two different closes by initial conditions separate exponentially in the course of time. The necessary requirements for a deterministic system to be chaotic are that the system must be nonlinear, and be at least three dimensional [12]. Each soft

computing technique can be used separately, but its complementary nature is its more powerful advantage. We can also create a hybrid system, combination of hard and soft computing, that can produce solutions to problems that are too complex or inherently noisy to tackle with conventional mathematical methods.

Potential Usages of Different Soft Computing Techniques in Software Reliability Models

Soft computing can be used for software faults diagnosis, reliability optimization and for time series prediction during the software reliability analysis. In this section we have discussed about application of soft computing technologies in software reliability.

Neural networks

Neural networks are simplified model of the biologic neuron system, it is massively parallel distributed processing system made up of highly interconnected neural computing elements that have the ability to learn and thereby acquire knowledge and make it available for use. Neural network has been applied for parameters estimation of the formal model and self learning process in order to predict the future outcomes. It has been shown that feed forward network can be applied for prediction. Back-error propagation is one of the most widely used neural network paradigms and has been applied successfully in a broad range of areas [13]. Karunanithi et al. [14,15] predict cumulative number of failure by design first neural network based software reliability model. They used feed-forward neural network, recurrent neural network and Elman neural network in their study and use execution time as the input of the network. They found that their models are better prediction models than some other statistical models [16]. Used connectionist models for software reliability prediction. Design the architecture of neural network by Falman's cascade correlation algorithm. They found that for end point prediction connectionist approach. Khoshgoftaar et al. [17] used the neural network for predicting the number of faults and introduced an approach for static reliability modeling. Then trained two neural networks; one with the complete set of principal components and one with the set of components selected by multiple regression model selection. Comparison of these models showed a better understanding of neural network software quality models. Sitte [18] compared, purposed neural network based software reliability prediction model, with recalibration for parametric models using some meaningful predictive measures with same datasets. Result showed that prediction with the help of neural network approach is better than others. Cai et al. [19] presented neural network based method for software reliability predictions, used back propagation algorithm for training. Performance of this purposed approach is evaluated by using different number of input nodes and hidden nodes. Result showed that its performance depends upon the nature of the handled data sets. Ho et al. [20] investigated a modified Elman recurrent neural network in modeling and predicting software failures and then performed a comprehensive study of connectionist models and their applicability to software reliability prediction and found them to be better and more flexible than the traditional models. Tian and Noore [21] proposed an on-line adaptive software reliability prediction model using evolutionary connectionist approach based on multiple-delayed-input single-output architecture, which showed better performance with respect to next-step predictability compared to existing NN model. Tian and Noore [22] presented an evolutionary neural network based method for software reliability prediction, used multiple-delayed-input single output architecture. Result showed that neural network architecture has a great impact on the performance of the network. Yu Shen Su et al. [23] purposed a model that uses the neural network approach to build a dynamic weighted combinational model. Then compared the performances of the neural network models with some conventional SRGMs from three aspects: goodness of fit, prediction ability for short-term prediction and long-term prediction. Result shows that purposed model has more accuracy with both goodness of fit and the prediction ability compared to existing conventional models. Viswanath [24] proposed two models such as neural network based exponential encoding and neural network based logarithmic encoding for prediction of cumulative number of failures in software. He use execution time as the input and applied on four data sets. Result showed that its result is better that other statistical model. Hu et al. [25] proposed an artificial neural network model to improve the early reliability prediction for current projects/releases by reusing the failure data from past projects/releases. Better prediction performance is observed in early phase of testing compared with original ANN model without failure data reuse. Su and Huang [26] proposed an artificial neural-network-based approach for software reliability estimation and modeling then further use of the neural network approach to build a dynamic weighted combinational model. The results obtained from the experiments show that the proposed model has a fairly accurate prediction capability

Kanmani et al. [27] purposed two neural network based software fault prediction models using Object-Oriented metrics. The results are compared with two statistical models using five quality attributes and found that neural networks do better Aljahdali et al. [28] explored connectionist artificial neural networks models as an alternative approach to derive these models by investigating the performance analysis of four different connectionist paradigms, multi-layer perceptron neural network, radial-basis functions, Elman recurrent neural networks and a neuro-fuzzy model, for modeling the software reliability prediction. In [29,30] Singh et al. used feed forward neural network for software reliability prediction and applied back propagation algorithm to predict software reliability growth trend. Also demonstrated a comparative analysis between the proposed approach and three well known software reliability growth prediction models using seven different failure datasets collected from standard software projects to test the validity of the presented method. Nirvikar Katiyar et al. [31] develop a non-parametric software reliability prediction system based on the neural network effects to improve the predictability by utilizing the diversity among the combined component neural networks. Huang et al. [32] derived software reliability growth models (SRGM) based on non-homogeneous poison processes (NHPP) using a unified theory by incorporating the concept of multiple change-points into software reliability modeling. Manjubala Bisi et al. [33] proposed a neural network based software reliability model to predict the cumulative number of failures based on Feed Forward architecture. The effect of encoding and the effect of different encoding parameter on prediction accuracy have been studied and its performance is tested using eighteen software failure data sets. Sandeep Kumar Jain et al. [34] proposed a method to estimate the reliability of the software consisting of components by using different neural network architectures. Then estimate the faults prediction behavior in the set of components over a cumulative execution time interval besides this the prediction of faults is estimated for the complete software. To predict the faults in each component of the software with the prediction of faults for the complete software for given cumulative execution time, apply the feed forward neural network architectures and its generalization capability.

Fuzzy system

Fuzzy Logic is derived from fuzzy set theory dealing with reasoning that is appropriate rather than precisely deduced from

classical predicate logic. A fuzzy model is a mapping between linguistic terms, attached to variables. Therefore the input to and output from a fuzzy model can be either numerical or linguistic [35]. Cai et al. [36] discussed the development of fuzzy software reliability models in place of probabilistic software reliability models (PSRMs). It was based on the proof that software reliability is fuzzy in nature. A demonstration of how to develop a fuzzy model to characterize software reliability was also presented. Khalaf khatatneh [37] developed a software reliability prediction model that implemented using the fuzzy logic technique. This model focused on a particular dataset behavior in predicting reliability. Focusing on a particular dataset behavior is performed to develop an accurate model since the recent work focused on developing a model which can be more accurate. Reformat [38] proposed an approach leading to a multi technique knowledge extraction and development of a comprehensive meta-model prediction system in the area of corrective maintenance of software. The system was based on evidence theory and a number of fuzzy-based models. Sultan Aljahdali et al. [39] investigated the use of fuzzy logic on building SRGM to estimate the expected software faults during testing process. Purposed model consists of a collection of linear sub-models, based on the Takagi-Sugeno technique and attached efficiently using fuzzy membership functions to represent the expected software faults as a function of historical measured faults. This purposed model gives a high performance modeling capabilities. S. Chatterjee et al. [40] two fuzzy time series based software reliability models have been proposed. The first one predicts the time between failures of software and the second one predicts the number of errors present in software. The purposed models are flexible, assumption free and very simple in computation. It not required any de-fuzzication techniques separately, which results in a significant reduction of computation time.

Neuro-fuzzy system

In neuro-fuzzy system we combine fuzzy logic and neural networks. It can be used for software reliability modeling investigation. Neuro-fuzzy describes a methodology for controlling neural networks by fuzzy logic. Kirti Tyagi et al. [41] proposed a model for estimating CBSS reliability, known as an adaptive neuro fuzzy inference system (ANFIS) that is based on these two basic elements of soft computing, neural network and fuzzy logic. ANFIS model gives a more accurate measure of reliability than the FIS model, as it reduces error from 11.74%, in case of FIS model, to 6.66% in ANFIS.

Genetic Algorithm (GA)

Genetic algorithm is a model of machine learning which derives its behavior from a metaphor of the process of evolution in nature. This is done by the creation within a machine of a population of individuals represented by chromosomes.

The fitness of each chromosome is determined by evaluating it against an objective function. To simulate the natural survival of the fittest process, best chromosomes exchange information to produce offspring chromosomes. The offspring solutions are then evaluated and used to evolve the population if they provide better solutions than weak populations members. Usually, the process is continued for a large number of generations to obtain a best-fit solution. Liang et al. [42,43] proposed a genetic algorithm optimizing the number of delayed input neurons and the number of neurons in the hidden layer of the neural network, predicting software reliability and software failure time. Sultan H. Aljahdali et al. [44,45] explored Genetic Algorithms (GA) as an alternative approach to derive software reliability models. For purposed system applicability three study sets; Military, Real Time Control and Operating System were used. Satya Prasad et al. [46] incorporate both imperfect debugging and change-point problem into the software reliability growth model (SRGM) based on the well-known exponential distribution the parameter estimation is studied. Proposed model is rated as better than the other considered models with respect to all the conditions are chosen.

Genetic Programming (GP)

Genetic programming can be viewed as an extension of the genetic algorithm, a model for testing and selecting the best choice among a set of results, each represented by a string. Genetic programming goes a step farther and makes the program or "function" the unit that is tested. Two approaches are used to select the successful program - cross-breeding and the tournament or competition approach. A difficult part of using genetic programming is determining the fitness function, the degree to which a program is helping to arrive at the desired goal. Costa et al. [47] Modeling software reliability growth with genetic programming. Experiments conducted to confirm the purposed hypothesis and demonstrate better results as compared to other traditional models and neural network model. E. Oliveira et al. [48] improve software reliability models using boosting techniques based on genetic programming. Boosting Technique combines several hypotheses of the training set to get better results. The most important improvement of this work is when consider models based on time and got excellent results by using just one function set. Y. Zhang et al. [49] Predicting for MTBF failure data series of software reliability by genetic programming algorithm. The evolution model of GP is then analyzed and appraised according to five characteristics criteria for some common-used software testing cases. Result showed the higher prediction precision and better applicability. Wasif Afzal et al. [50] discussed the suitability of using GP for software reliability growth modeling and highlight the mechanisms that enable GP to progressively search for fitter solutions. The experiments of using GP for software reliability growth modeling have indicated positive results. Eduardo Oliveira Costa et al. [51] introduced a new GP based approach, named (μ+λ)GP. This algorithm was introduced to improve the performance of GP. To evaluate this purposed algorithm, two kinds of models: based on time and on coverage were presented for experimental results, which is always better than classical GP. Zainab Al-Rahamneh et al. [52] proposed the use of Genetic Programming (GP) as an evolutionary com-putation approach to handle the software reliability modeling problem. Evaluate the GP developed model and results show that this purposed model is superior to other models such as Yamada S-shaped, Generalized Poisson, NHHP and Schneidewind reliability models.

Artificial Bee Colony (ABC)

Dervis Karaboga, in 2005, defined a new algorithm, motivated by the intelligent behavior of honey bees known as artificial bee colony. It is an optimized tool provides a population-based search procedure in which individuals called foods positions are modified by the artificial bees with time and bee's aim to discover the places of food sources with high nectar amount and finally the one with highest nectar. Tarun Kumar Sharma et al. [53] proposed a modified version of the ABC, the DABC (Dichotomous ABC), to improve its performance, in terms of converging to individual optimal point and to compensate the limited amount of search moves of original ABC. Also explored the applicability of the modified artificial bee colony algorithm to estimate the parameters of software reliability growth models (SRGM). The estimated model parameters were used to predict the faults in a software system during the testing process.

Ant colony

Ant Colony Optimization [54] is a technique which uses probability to solve problems where the computations are reduced with the help of

graphs to get efficient paths. It has been applied to many fields as its robustness and is easy to collaborate with other methods. It has well performance to the optimization problem and has a good convergence rate. Changyou Zhenga et al. [55] proposed a parameter estimation method based on the Ant Colony Algorithm. Its results are derived from Numerical examples based on three sets of real failure data. Purposed method shows higher precision and faster convergence speed. Latha Shanmugam et al. [56] discussed a parameter estimation method based on Ant Colony Algorithm. The outcome of the experiment using six typical models demonstrated that this algorithm can be applied for estimating parameters. Higher precision and faster convergence speed is achieved through this method when compared with PSO algorithm. Latha Shanmugam et al. [57] studied enhancement and Comparison of Ant Colony Optimization Methods for Software Reliability Models. The Enhanced method shows significant advantages in finding the goodness of fit for software reliability model such as finite and infinite failure Poisson model and binomial models. It is comparatively giving better Estimation Accuracy (approximately 10%) than the existing Ant Colony Optimization.

Simulated Annealing (SA) algorithm

Simulated annealing (SA) is an iterative search method inspired by the annealing of metals [58,59] Starting with an initial solution and armed with adequate perturbation and evaluation functions, the algorithm performs a stochastic partial search of the state space. Nidhi Gupta et al. [60] the simulated annealing technique of mean field approximation for finding the possible minimum number of failed components in the sequential testing. These minimum numbers of failed components are depending upon the selection of time intervals or slots. Also purposed a new energy function with the mean field approximation. The algorithm of the whole process shows that this approach can generate the optimal solution. Pai and Hong [61] applied support vector machines (SVMs) for forecasting software reliability where simulated annealing (SA) algorithm was used to select the parameters of the SVM model. The experimental results reveal that the SVM model with simulated annealing algorithms (SVMSA) results in better predictions than the other methods. Mohamed Benaddy et al. [62] presented a hybrid approach based on the Neural Networks and Simulated Annealing. An adaptive simulated Annealing algorithm is used to optimize the mean square of the error produced by training the neural network, predicting software cumulative failure. The purposed adaptive Simulated Annealing gives better performance in execution time than the Real Coded Genetic Algorithm (RCGA), because of the search space, which reduced from a population of solutions for the RCGA to one solution for the proposed Simulated Annealing.

Tabu search algorithm

The Tabu Search (TS) is an optimization method, based on the premise that problem solving, in order to qualify as intelligent, and must incorporate adaptive memory and responsive exploration [63]. The Tabu method was partly motivated by the observation that human behavior appears to operate with a random element that leads to inconsistent behavior given similar circumstances. Caserta et al. [64] presented a new meta- heuristic-based algorithm for complex reliability problems. The algorithm effectively uses features of the Tabu Search paradigm, with special emphasis on the exploitation of memory-based mechanisms. It balances intensification with diversification via the use of short-term and long-term memory. The proposed algorithm proves to be robust with respect to its parameters and it is especially suited for very large scale instances of the reliability problem, when exact approaches are doomed to fail.

Cuckoo search algorithm

This algorithm is based on the obligate brood parasitic behavior of some cuckoo species in combination with the L´evy flight behavior of some birds and fruit flies. Cuckoo search algorithm is very successful in finding good and acceptable solutions to the problem of parameter estimation of Software Reliability Growth Models. This algorithm search strategy can efficiently navigate throughout the search space of the problem and locate very good solutions using fewer iterations and smaller populations. Najla Akram AL-Saati et al. [65] estimated parameters based on the available failure data. Cuckoo Search outperformed both PSO and ACO in finding better parameters tested using identical datasets, but worse in case of extended ACO. The Exponential, Power, S-Shaped, and M-O models are considered in this work. The search strategy of the cuckoo can efficiently navigate throughout the search space of the problem and locate very good solutions using fewer iterations and smaller populations

Summarization

In the past few decades a number of software reliability models have been analyzed, designed and evaluated. Soft computing plays an important role in the recent advancements in the software reliability growth models. Today these models included the application of different soft computing techniques such as Neural Network (NN), Fuzzy Logic, Genetic Algorithms (GA), Genetic Programming (GP), Artificial Bee Colony (ABC) and Ant Colony etc., a brief summary of soft computing techniques in software reliability models have been summarized in this section as shown in Table 1.

In this summary, we notice that different soft computing techniques are used in different shapes with these models. We observed that Neural Network approach is more liked by the researchers in software reliability models. Genetic Programming provides more accuracy than other soft computing techniques. Cuckoo Search, Stimulated Annealing and Tabu Search are used in this field but not so widely yet. This table data is important in case of comparison and selection of soft computing technique in terms of modeling capabilities.

Comparison of Different Soft Computing Techniques in Terms of Software Reliability Models

Comparisons is very useful in case of optimal selection, user can view all possible choices on a single plate form and can select the best suited as his/her requirement. In Table 2, we have compared different soft computing techniques in terms of software reliability modeling capabilities such as data sets, re-adjustments for new data set, process visibility, facts and outputs etc. This comparison has outlined some parameters of modeling capabilities. From this table we observed that all the techniques explain its outputs and are applicable for complex models except genetic algorithm. Comparison revealed that only fuzzy Logic can be widely used for all the modeling capabilities. Ant colony, Stimulated Annealing, Tabu Search and Cuckoo Search can also be used for most of the modeling capabilities except only small data set capabilities. The rapid growth of soft computing techniques suggests that the impact of these algorithms will be used increasingly for software reliability models in the coming years. This table will help computer scientist who are keen to contribute their works to the field of software reliability.

Conclusion

In this paper, we have discussed about the work done by the various researchers, with the endeavor made to include as many references as possible from year 1990 to 2014. Based on this paper, we thrash out some

s.no	Study	Technology used	Project data	Summary of Result	References
1	N.Karunanithi,Y.K.Malaiya and D. Whitley	Neural Network	DS-1K.Matsumoto(1988) DS-2J.D. Musa(1987)DS-3M. Ohba(1984)	Explores the use of feed-forward neural networks as a model for software reliability growth prediction	[14]
2	N.Karunanithi,D.Whitley, Y.K. Malaiya	Neural Network	DS form Yoshiro Tohma project	After training the neural network with afailure history up to time t you can use thenetwork to predict the cumulative faults atthe end of a future testing and debuggingsession.	[15]
3	Taghi M.Khoshgoftaa,RobertM. Szabo and Peter J.Guasti	Neural Network	Data collected from three similar systems(Kernel2,Kernel3andKernel4)and twodissimilarsystems(Kernel5andKernel6)	The excellent predictive results observed inthe neural network models indicate thatneural networks should be seriouslyconsidered as an effective modeling too forsoftware engineers	[17]
4	Renate Sitte	Neural Network	Musa data-sets S1 & SS3	Neural Networks are not only much simpler to use than the recalibration method, but that they are equal or better trend predictors	[18]
5	S.L.Ho,M.Xie andT.N.Goh	Neural Network	DS-1:Military computer system DS-2:Musa-OkumotoModelDS-3: Goel Okumoto Model	The Elman model is comparatively better than the Jordan model and very much superior than the feed -forward model	[20]
6	NirvikarKatiyarand Raghuraj Singh	Neural Network	DS1and DS2 from handbookofsoftwarereliability engg.LyuM.R.(1996)New York	Purposed system achieves significantly lower prediction error compared with the single NN and traditional SRGMs	[31]
7	Sultan Aljahdali et al.	Neural Network	Real-time command and Control processing Commericial and military applications	NN provide models with smaller normalized root of mean of the square of error than the regression model in all considered cases	[]
8	Yu Shen Su, Chin-Yu Huang,YiShin and Jing Xun Chen	Neural Network	Realcommand & control project, John D Musa, Bell Lab	Achieve a dynamic weighted combinational model	[23]
9	Sultan H. Aljahdali and Khalid A. Buragga	Neural Network	Rea l time control project	The Elman recurrent NN is a robust technique for function prediction due capturing the dynamic behavior of the data set.	[28]
10	Manjubala Bisi et al.	Neural Network	18differentdatasets(Military, Real time System,RealtimeCommandandControl,On-line data Entry etc.)	A Feed Forward neural network with two encoding scheme such as exponential andlogarithmic function has been proposed.	[33]
11	Sandeep Kumar Jain & Manu Pratap Singh	Neural Network	Data collected from localtraining set.	Estimated the faults prediction behavior in the set of components over a cumulative execution time interval besides this the prediction of faults is estimated for the complete software.	[34]
12	S. Chatterjee, S. Nigam,Singh,Upadhyaya	Fuzy Logic	DS-1: Musa J D(1975) software reliability data DS-2: Pham H(2006) system software reliability	Purposed models are flexible, assumption free and very simple in computation. It not requiredanyde-fuzzifiedtechniques separately and computation time is reduced	[40]
13	Sultan Aljahdali	Fuzy Logic	Realcommand & control project, Militaryandoperatingsystem,JohnDMusa, Bell lab.	Developed models provide high performance modeling capabilities.	[39]
14	Khalaf Khatatneh andThaer Mustafa	Fuzy Logic	datasetfrom command and controlapplications,Musa,John D.	Developed model can predict accurate results in most points of the target database	[37]
15	Sultan H. Aljahdali and MohammedE.El-Telbany	Genetic Algorithm	Data from threeprojects Military,RealSystem Controland Operating system	Measured the predictability of software reliability using ensemble of models which performed better than the single model and also find that the weighted average combining method for ensemble has a better	[44]
16	Sona Ahuja, Guru Sarand Mishra and Agam Prasad Tyagi	Genetic Algorithm	Musa Data-set, software reliability prediction andapplication, 1985	GA based hybrid stochastic search technique,has turned out to be good tool for optimizedsimulated trajectory for variable which areimportant performance indicatorto predictqualityof reliability of the predicted softwarefailure	[66]
17	Sultan H. Aljahdali and Mohammed E.	Genetic Algorithm	Data from three projects.They are Military,Real Time Control and Operating System.	As far as the predictability of the single AR model and ensemble of AR models trained byGA algorithm over the trained and test data isconcerned,the ensemble of modelsperformed better than the single model. Also,find that the weighted average combiningmethod forensemble has a betterperformance in a comparison with averagemethod.	[45]
18	R.Satya Prasad, O.Naga Raju and R.R.L Kantam	Genetic Algorithm	DS-1:Misra,P.N.,1983Software reliability analysis.IBM Syst. DS-2:Pham,H.,1993.Software reliability assessment:Imperfect debugging and multiple failure types in softwaredevelopment	This model is rated as better than the other considers models with respect to all conditions are chosen.	[46]
19	Eduardo Oliveira Costa et al.	Genetic Programming	DS-1John Musa at Bell Telephone Laboratories DS-2 failure data of a program called Space	GP is a suitable tool to discover an equation to modeling software reliability and is also able to discover the equation that better represents the data [	[47]

20	E.Oliveira,A.Pozo, and S.Vergilio	Genetic Programming	DS-1 John Musa atBellTelephone LaboratoriesDS-2failure data of a program called Space	Use boosting techniques to improve software reliability models based on Genetic Programming.	[48]
21	Wasif Afzal and Richard Torkar	Genetic Programming	Data available at:http://www.gp-field-guide.org.uk,2008.	The experiments of using GP for software reliability growth modeling have indicated positive results, which warrant further investigation with larger real-world industrial data sets.	[50]
22	Eduardo Oliveira Costa, Aurora Trinidad Ramirez Pozo, and Silvia Regina Vergilio	Genetic Programming	DS-1:software reliability data by J Musa(1980)DS-2: space program	Purposed(μ+λ) GPsystem results arealways better thanthe classical technique results also improves the performance forsmall datasets	
23	Zainab Al-Rahamneh, Mohammad Alaa F. Sheta, Sulieman Bani-Ahmad, Saleh Al-Oqeili	Genetic Programming	Data set from Y. Tohmapublished in IEEEtransactions onSoftwareEngineering, Vol. 15, No. 3,1989	Here adopted recalibrated and adjusted GP operators to speed up the convergence process	[51]
24	Tarun Kumar Sharma, MilliePant and Ajith Abraham	Artificial Bee Colony	IEEE Congress onEvolutionary Computation,Sheraton,VancouverWallCentre,Vancouver,BC,Canada,2006.,10428–10435	Modified version have a better success rate than original ABC	[52]
25	Changyou Zhenga,Xiaoming Liua,Song Huanga and Yi Yaoa	Ant Colony	SYS1, SYS2, SYS3 from Musa dataset	Experiments with three typical models show that this ant colony based algorithm demonstrates good applicability	[53]
26	Latha Shanmugam and Lilly Florence	Ant Colony	Musa real time data-set	Based on the results, it was found that the Enhanced Ant Colony Optimization Method is giving better estimation accuracy than Existing ACO method. Time and Space Complexity is also reduced.	[55]
27	The UK's expert	Ant Colony	failure data which is given in the Musa data Set from the DACS Web Site	Based on the results, we found that the proposedSimulatedAnt Colony Optimization Method is giving 15% better estimation accuracy	[57]
28	Najla Akram AL-Saati	Cuckoo Search	first group dataset: A Sheta. Et al. (2006)second group dataset:John Musa of Telephone Laboratories	CS do better than both PSO and ACO infinding better parameters tested using identical datasets.	[65]
29	Nidhi Gupta and Manu Pratap Singh	Stimulated Annealing	a series of tests conducted under certain stipulated conditions on 1,000 software components	Results show that after applying the sequential testing with MFA we can optimizethe number of failures up to a minimumvalue	[60]
30	Mohamed Benaddy and Mohamed Wakrim	Stimulated Annealing	John Musa of Bell Telephone Laboratories	The Performance in execution time of the proposed adaptive simulated Annealing is better than the RCGA, because of the search space, which reduced from a population of solutions for the RCGA to one solution for the proposed Simulated Annealing	[62]
31	Kirti Tyagi and Sharma	Neuro Fuzzy	collecte data from 47 classroom-based projects	Results show that the ANFIS improves the reliability evaluation of the FIS technique	[41]
32	M. Casertaa,and A.Márquez Uribe	Tabu Search	Berman and Ashrafi (1993)	Proposed algorithm is robust with respect to its parameters and it is especially suited for very large scale instances of the reliability problem.	[64]

Table 1: Summary of Soft Computing Techniques in Software Reliability Models.

Sr. No	Technology Used	Explain outputs	Suitability for small data sets	Can be re-designed for new data set	Reasoning process is visible	Applicability for complex models	Either known facts considered
1	Neural	No	No	No	No	Yes	Partially
	Networks						
2	Fuzzy Logic	Yes	Yes	Yes	Yes	Yes	Yes
3	Genetic	Partially	Partially	Yes	Yes	Partially	No
	Algorithms						
4	Genetic	Yes	No	No	No	Yes	No
	Programming						
5	Artificial Bee	Yes	Partially	Partially	No	Yes	Yes
	Colony						
6	Ant Colony	Yes	No	Yes	Yes	Yes	Yes
7	Stimulted	Yes	No	Yes	Yes	Yes	Yes
	Anealing						
8	Tabu Search	No	Partially	Yes	Yes	Yes	Yes
9	Cuckoo Search	Yes	No	Yes	Yes	Yes	Yes

Table 2: Comparison of Soft Computing Techniques in terms of Modeling Capabilities.

soft computing techniques, such as: Neural networks (NN), Fuzzy Logic (FL), Genetic Algorithm (GA), Genetic Programming (GP), Artificial Bee Colony (ABC) and Ant Colony etc. We emphasized on the role of existing soft computing techniques in software reliability modeling, with the reliance that it would serve as a reference to both old and new, incoming researchers in this field, to support their understanding of current trends and assist their future research prospects and directions. Further we compared soft computing techniques in terms of modeling capabilities, which enhances the selection process of soft computing technique for software reliability models.

References

1. Agarwal kk, Singh Y (2005) Software Engineering. New Age International Publisher, New Delhi.
2. Kapur PK, Garg RB (1990) Cost reliability optimum release policies for a software system with testing effort. Operations Research 27: 109-116.
3. Musa JD (1999) Software Reliability Engineering: More Reliable Software, Faster Development and Testing, McGraw-Hill.
4. Musa JD, Iannino A, Komodo K (1987) Software Reliability: Measurement, Prediction and Application. McGraw-Hill.
5. Bonissone P, Chen YT, Goebel K, Khedkar P (1999) Hybrid Soft Computing Systems: Industrial and Commercial Applications. Proceedings of the IEEE 87: 1641-1667.
6. Das SK, Kumar A, Das B, Burnwal AP (2013) On Soft computing Techniques in Various Areas. ACER.
7. Zadeh LA (1968) Fuzzy Algorithms. Information and Control 12: 94-102.
8. Zadeh LA (1965) Fuzzy Sets. Information and Control 8: 338-353.
9. Kuo W, Prasad VR (2000) An annotated Overview of System Reliability Optimization. IEEE Transaction on Reliability 49: 176-186.
10. Streichert F (2002) Introduction to Evolutionary Algorithms. Frankfurt MathFinance Workshop, University of Tubingen, Germany.
11. Abraham A (2005) Handbook of Measuring System Design. In: Peter H, Sydenham, Thorn R (eds.). Evolutionary Computation. John Wiley & Sons, Ltd.
12. Boccaletti S (2000) The control of chaos: theory and applications. Elsevier Science 329: 103-197.
13. Aggarwal G, Gupta VK (2013) Neural Network Approach to Measure Reliability of Software Modules: A Review. International Journal of Advances in Engineering Sciences.
14. Karunanithi N, Malaiya YK, Whitley D (1992) Prediction of software reliability using neural networks. Proceedings of the Second IEEE International Symposium on Software Reliability Engineering.
15. Karunanithi N, Whitley D, Malaiya YK (1992) Using neural networks in reliability prediction. IEEE Software 9: 53-59.
16. Karunanithi N, Whitley D, Malaiya YK (1992) Prediction of software reliability using connectionist models. IEEE Trans Software Engg 18: 563-573.
17. Khoshgoftaar TM, Szabo RM, Guasti PJ (1995) Exploring the Behavior of Neural-network Software Quality Models. Software Engg J 10: 89-96.
18. Sitte R (1999) Comparison of Software Reliability Growth Predictions: Neural networks vs. parametric recalibration. IEEE transactions on Reliability 48: 285-291.
19. Cai KY, Cai L, Wang WD, Yu ZY, Zhang D (2001) On the neural network approach in software reliability modeling. Journal of Systems and Software 58: 47-62.
20. Ho SL, Xie M, Goh TN (2003) A Study of the Connectionist Models for Software Reliability Prediction. Computers and Mathematics with Applications 46: 1037-1045.
21. Tian L, Noore A (2005) On-line prediction of software reliability using an evolutionary connectionist model. Journal of Systems and Software 77: 173-180.
22. Tian L, Noore A (2005) Evolutionary neural network modeling for software cumulative failure time prediction. Reliability Engineering and System Safety 87: 45-51.
23. Su SY, Huang CY, Chen YS, Chen JX (2005) An Artificial Neural-Network-Based Approach to Software Reliability Assessment. TENCON, IEEE Region 10: 1-6.
24. Viswanath SPK (2006) Software Reliability Prediction using Neural Networks. PhD. Thesis, Indian Institute of Technology Kharagpur.
25. Hu QP, Dai YS, Xie M, Ng SN (2006) Early software reliability prediction with extended ANN model (2006) Proceedings of the 30th Annual International Computer Software and Applications Conference, pp. 234-239, 2006.
26. Su YS, Huang CY (2006) Neural-network-based approaches for software reliability estimation using dynamic weighted combinational models. Journal of Systems and Software 80: 606–615.
27. Kanmani S, Uthariara VR, Sankaranarayanan V, Thambidurai P (2007) Object-oriented software failure fault prediction using neural networks. Information and Software Technology 49: 483-492.
28. Aljahdali AS, Buragga KB (2008) Employing four ANNs Paradigms for Software Reliability Prediction: an Analytical Study. ICGST-AIML Journal.
29. Singh Y, Kumar P (2010) Application of feed-forward networks for software reliability prediction. ACM SIGSOFT Software Engineering Notes 35: 1-6.
30. Singh Y, Kumar P (2010) Redirection of Software Reliability Using feed Forward Neural Networks. International conference on Computational Intelligence and software Engineering.
31. Katiyar N, Singh R (2011) Effect of Neural Network for Prediction of Software Reliability. VSRD-IJCSIT 1: 490-500.
32. Huang CY, Lyu MR (2011) Estimation and Analysis of Some Generalized Multiple Change-Point Software Reliability Models. IEEE Transaction on Reliability 60: 498-514.
33. Bisi M, Goyal NK (2012) Software Reliability Prediction using Neural Network with Encoded Input. International Journal of Computer Applications.
34. Jain SK, Singh MP (2013) Estimation for Faults Prediction from Component Based Software Design using Feed Forward Neural Networks. IJARCCE.
35. Aljahdali S, Debnath NC (2004) Improved Software Reliability Prediction through Fuzzy Logic Modeling. IASSE.
36. Cai KY, Wen CY, Zhang ML (1991) A critical review on software reliability modeling. Reliability Engineering and System Safety 32: 357-371.
37. Khatatneh K, Mustafa T (2009) Software Reliability Modeling Using Soft Computing Technique. European Journal of Scientific Research 26: 147-152.
38. Reformat M (2005) A fuzzy-based multimodel system for reasoning about the number of software defects. International Journal of Intelligent Systems 20: 1093-1115.
39. Aljahdali S (2011) Development of Software Reliability Growth Models for Industrial Applications Using Fuzzy Logic. Journal of Computer Science 7: 1574-1580.
40. Chatterjee S, Nigam S, Singh JB, Upadhyaya LN (2011) Application of Fuzzy Time Series in Prediction of Time Between Failures & Faults in Software Reliability Assessment. Fuzzy Inf Eng 3: 293-309.
41. Tyagi K, Sharma A (2014) An adaptive neuro fuzzy model for estimating the reliability of component-based software systems. Applied Computing and Informatics 10: 38-51.
42. Tian L, Noore A (2005) Evolutionary neural network modeling for software cumulative failure time prediction. Reliability Engineering & System Safety 87: 45-51.
43. Tian L, Noore A (2005) On-line prediction of software reliability using an evolutionary connectionist model. Journal of Systems and Software 77: 173-180.
44. Aljahdali SH, El-Telbany ME (2008) Genetic Algorithms for Optimizing Ensemble of Models in Software Reliability Prediction. ICGST-AIML.
45. Aljahdali SH, El-telbany ME (2009) Software Reliability Prediction Using Multi-Objective Genetic Algorithm. IEEE.
46. Prasad RS, NagaRaju O, Kantam RRL (2010) SRGM with Imperfect Debugging by Genetic Algorithms. International Journal of software engineering & applications 11.
47. Costa E, Vergilio S, Pozo A, Souza P (2005) Modeling software reliability growth with genetic programming. ISSRE.

48. Oliveira E, Pozo A, Vergilio S, (2006) Using boosting techniques to improve software reliability models based on genetic programming. ICTAI.

49. Zhang Y, Chen H (2006) Predicting for MTBF failure data series of software reliability by genetic programming algorithm. In Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications, Washington, USA.

50. Afzal W, Torkar R (2008) Suitability of genetic programming for software reliability growth modeling. IEEE

51. Costa OE, Pozo ATR, Vergilio SR (2010) A Genetic Programming Approach for Software Reliability Modeling. IEEE Transactions on Reliability.

52. Al-Rahamneh Z, Reyalat M, Sheta FA, Bani-Ahmad S, Al-Oqeili S (2011) A New Software Reliability Growth Model:Genetic-Programming-Based approach. Journal of Software Engineering and Applications, 4: 476-481.

53. Sharma TK, Pant M, Abraham A (2011) Dichotomous Search in ABC and its Application in Parameter Estimation of Software Reliability Growth Models. IEEE.

54. Li, W.,Q. Yin and X. Zhang (2010) Continuous quantum ant colony optimization and its application to optimization and analysis of induction motor structure. Proceedings of the IEEE 5th International Conference on Bio-Inspired Computing: Theories and Applications, Changsha.

55. Zhenga C, Liua X, Huanga S, Yaoa Y (2011) A Parameter Estimation Method for Software Reliability Models. Procedia Engg 15: 3477-3481.

56. Shanmugam L, Florence L (2012) A Comparison of Parameter best Estimation Method for software reliability models. International Journal of Software Engineering & Applications.

57. Shanmugam L, Florence L(2013) Enhancement and comparison of ant colony optimization for software reliability models. Journal of Computer Science 9: 1232-1240.

58. Kirkpatrick S, Gelatt C, Vecchi M (1983) Optimization by simulated annealing. Science, 220: 498– 516.

59. Cerny V (1985) Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. Journal of Optimization Theory and Application 45: 41–51.

60. Gupta N, Singh MP (2006) Evolutionary algorithms, simulated annealing and tabu search: a comparative study. IJE Transactions B: Applications.

61. Pai PF, Hong WC (2006) Software reliability forecasting by support vector machines with simulated vector machines with simulated annealing algorithms. The Journal of Systems and Software 79: 747–755.

62. Benaddy M, Wakrim M (2012) Simulated Annealing Neural Network for Software Failure Prediction. International Journal of Software Engineering and Its Applications.

63. Glover F, Laguna M (1997) Tabu Search: A Tutorial. Kluwer Academic Publishers, Boston.

64. Caserta M, Uribe MA (2006) Tabu search-based meta-heuristic algorithm for software system reliability problems. Computers & Operation Research 36: 811-822.

65. Akram N. AL-Saati, Abd-AlKareem M (2013) The Use of Cuckoo Search in Estimating the Parameters of Software Reliability Growth Models. International Journal of Computer Science and Information Security.

66. Ahuja S, Mishra GS, Tyagi AP (2002) Jelinski – Moranda Model for Software Reliability Prediction and its G. A. based Optimised Simulation Trajectory.D. E. I. Dayalbagh, Agra.

Citation: Kaswan KS, Choudhary S, Sharma K (2015) Software Reliability Modeling using Soft Computing Techniques: Critical Review. J Inform Tech Softw Eng 5: 144. doi:10.4172/2165-7866.1000144

Submit your next manuscript and get advantages of OMICS Group submissions

Unique features:

- User friendly/feasible website-translation of your paper to 50 world's leading languages
- Audio Version of published paper
- Digital articles to share and explore

Special features:

- 400 Open Access Journals
- 30,000 editorial team
- 21 days rapid review process
- Quality and quick editorial, review and publication processing
- Indexing at PubMed (partial), Scopus, EBSCO, Index Copernicus and Google Scholar etc
- Sharing Option: Social Networking Enabled
- Authors, Reviewers and Editors rewarded with online Scientific Credits
- Better discount for your subsequent articles

Submit your manuscript at: http://www.omicsonline.org/submission/

International Journal of Sustainable Energy and Environmental Research

journal homepage: http://www.pakinsight.com/?ic=journal&journal=13

SUSTAINABLE HANDLING OF CONSTRUCTION AND DEMOLITION (C & D) WASTE

Shishir Bansal[1]* --- S K Singh[2]

[1]*Research Scholar, Environmental Engineering Department, Delhi Technological University, Delhi, India*
[2]*Head, Environmental Engineering Department, Delhi Technological University, Delhi, India*

ABSTRACT

Development as a standalone feature has got no status, unless it is sustainable. Development may include active actions in various sectors like medical field, information technology, education etc. But, one field that is an inherent part of all is the construction sector. With the development of society and for the development of society at all fronts, construction activities are seen everywhere. Along with the construction activities, the demolitions of existing structures, which have either outlived their service life or otherwise needing replacement, have to be demolished. It is estimated that the construction industry in India generates about 10-12 million T of waste annually. The waste generation in Delhi itself is estimated to be around 5000 T per day. Thus a need is there to find solution for its sustainable disposable. Dumping of C&D waste is not only unauthorized, but also anti-environmental. It is a huge challenge to the mankind to find sustainable solutions for safe and secure reuse or recycling, so that nothing is required to be disposed or say dumping in an unauthorized manner. There is also a huge demand for natural aggregates in the construction sector with a significant gap in its demand and supply, which can also be reduced marginally by the recycling of construction and demolition waste. Proper handling, storage and treatment of C&D waste not only prevent degradation of Mother Earth, but also have significant impact on sustainability by way of reducing the use of natural resources. The paper covers various issues related to the reusing and recycling of C&D waste, required regulatory mechanism and procedures to be followed for achieving the aim with an ulterior motive of Environmental sustainability.

© 2015 Pak Publishing Group. All Rights Reserved.

Keywords: Construction & demolition, Reuse, Recycling, Technology, Equipment.

Contribution/ Originality

This study is one of very few studies, which have investigated into the effective methods and appropriate tools to carry out the demolition of existing structures so that the debris can be put into

**Corresponding Author*

the reuse or recycling operations, getting useful products and also save Mother Earth from unsustainable degradation.

1. INTRODUCTION

Our Society is growing in all sectors like creation of more medical facilities, more education facilities, more industries, more housing etc. Every growth involves heavy construction activities and the quantum is never such before in past. Construction activities and demolition activities are like sightseeing for a tourism industry. Both are going hand in hand. Whereas construction is due to rebuilding of old structures or new structures, demolition is growing due to old structures needs replacement with time and growing needs of the society. When the structures were created long back, the population was lesser and also the need of dwelling units was proportionately much lesser. Therefore lavish sized structures were created with more in area and lesser in numbers. Now, need is to create more number of units in the same land coverage for a better satisfaction of large sized population. When horizontal expansion is no more available, this is possible by reconstructing the house with vertical expansion and creating more number of such dwelling units. All such activities are generating huge amount of waste, called the Construction and Demolition waste. Disposal of such debris in a safe environment is a big challenge for the builders, developers and owners.

When on one hand the disposal of debris is a challenge, then on the other hand there is an acute shortage of naturally available aggregates for construction of buildings. Reduction of this demand is possible only with the reusing or recycling of waste generated from the construction activities. Hence, the construction sector must accept the use of C& D waste wherever feasible. This will also lead to a little saving in virgin raw material, but significantly contribute towards the reduction in waste disposal.

2. REUSING & RECYCLING: 2 PILLARS OF WASTE STREAM UTILIZATION (Dhir and Paine, 2010; Mullick, 2012)

2.1. Reusing of C& D Waste

Reusing of construction and Demolition waste is different from Recycling**.** It does not require any further processing to convert into a useful product. The items which are usable directly are screened out from the debris and put into the possible use without further processing or the application of any further energy for conversion into the useful product. For example, full bricks can be screened out of the demolition debris and used as a partition wall in a building. Otherwise same would have been converted into smaller pieces for using it as an aggregate or brick bats for plinth protection etc. Thus reusing a waste item is a better contribution to the environment as it is saved from further impacts due to recycling activities.

Since the reusing of C&D waste is always more advantageous, it is essential that to have more and more reusable materials in debris. This is possible, if sufficient precautions are taken while a building is demolished. There should be an engineered deconstruction plan instead of just converting the standing structure into debris within minutes. Useful products like doors and

windows, bricks, reinforcement, from RCC components, structural steel can be taken out with little extra efforts and put into reuse without much processing.

2.2. Recycling of C& D Waste

Once the waste generated from construction and demolition activities has been segregated and reusable items are taken out, the leftover is available for further processing i.e. recycling into next useful stage. Recycling of this waste into useful products to extend the service to environment is a challenge.

If we see the world wide, there is enough scope of recycling of such a waste, but at present, in India, the recycling of Construction and Demolition waste has started now only and we have to go miles in this direction. In New Delhi, only one pilot plant is functioning at Burari in North Delhi, but it requires many such plants to fulfill the needs in Delhi after seeing the recycling potential of waste generated daily in New Delhi. About 5000 MT of recyclable waste is generated daily in New Delhi.

2.3. Concept, Education and Sensitization

Since the concept of C& D waste management is new, it is essentially required to spread the education and sensitize the Engineers as well as Developers in real Estate. Unless public support is there, the purpose of actual and effective recycling and promoting the use of recycled products cannot be achieved. It is essentially required to spread the concept and spread the education in this field in order to sensitize all the stakeholders for sustainable promotion of C& D waste reusing and recycling. The present mindset of public and their attitude towards the waste generated from construction and demolition sites is required to be changed. It is required to sensitize not only the Engineers and stack holders, but also the regulatory authorities in construction industry. One must appreciate the potential of reusing and recycling of C& D waste and technical aspects involved in this this along with the machinery required for demolition in Engineering manner as well as recycling in an Engineering manner. This will enhance the effective use of recycled materials and deployment of best practice for implementation, enforcement and achieving the aim with an ulterior motive of Environmental sustainability.

Recycling cost is influenced by transportation distance and amount of waste concrete to be recycled. CO_2 emission is influenced by transportation distance. It is important to minimize C&D waste generation and maximize reuse/recycling as the construction industry is consumer of tremendous amount of natural resources and energy as well as emitter of GHGs. Establishment of effective strategies and enactment of laws and regulations is essential to achieve this. In addition, provision of some incentives to users of the recycled products deems to be necessary to promote the use. It is essential to assess the life-cycle as it provides quantitative tool to assess environmental impact of C&D waste reuse/recycling.

3. ENGINEERING THE DEMOLITION OF BUILDINGS

In order to make the full use of C&D waste, it is essential that the building which is to be reconstructed is brought down in an Engineering manner so that the demolition waste does not

become debris and left out for only disposal either on a dumping site or at most used for filling. Today, the demolition is not just a demolition which can be performed by any contractor in any manner suitable as per his whims and fancies. It is a specialized Engineering which is as precise as construction of a new building, if one is sensitive towards the related Environmental issues, social issues and economics related with the demolition.3

3.1. Environmental Issues Related with Demolition

A building should be demolished in such a manner that maximum reusable components are safely dragged from the building using suitable extractors, chippers or any other suitable machinery. After the reusable components are segregated, recyclable components should be screened out and sent for recycling plant whether it is simply concrete or bricks or even metals or plastic components. Only the non-reusable and non-recyclable components are to be dumped in a safe and sound manner that to the approved dumping stations.

3.2. Social Issues

Safe and sound demolition of a building requires attention towards safety issues of the workers deployed in demolition. In the absence of appropriate machinery, it is going to put the labors on the job to a big risk and it will not amount to extraction of best results as afar as re-using and recycling is concerned. Many of the reusable / recyclable components will be lost as debris leftovers just with a fate of getting dumped below the earth. Another important factor to look into this aspect that skilled labor for demolition will not be always available. In such a situation it is the need to go for the most appropriate tools and machinery as partial replacement of the manpower.

3.3. Economic Issues

One must understand the limitation of the work output of a labor in any activity. When demotion is the activity, it is not only the efficiency of labor that counts, but the final productivity also matters. Appropriate tools and machinery will always result in maximum extraction of usable components and minimize the debris to be dumped. Thus it results in some contribution to the cost of construction and also reduced the cost of safe disposal of debris.

4. TOOLS AND MACHINERY

There are safe systems and supporting machinery available so that the building is brought down in the planned safe manner and also full potential of reuse of demolished material is made use of. Useful reusable materials like wooden materials including doors and windows, metallic beams or bricks can be taken out and put into the appropriate use without putting them into recycling system.



Fig-1. Wrecking Ball: Uncontrolled demolition



Fig-2. Mounted Hydraulic Breakers

It has been observed that in the past when safe and sound system equipped with appropriate machineries was not available, wrecking ball was used to crush the building wall. It was totally uncontrolled mechanism and many a times fatal accidents have occurred at work site due to lack of control on its movement, transfer of energy to the existing structure and throw of debris in all possible directions. In order to control such demolitions, hydraulic excavators with suitable attachments can make the task easier and better outcomes as far as reusable materials are concerned. It is always appreciated, if more and more reusable materials are taken out as it does not require any further energy for recycling. Such attachments may include hydraulic breakers, demolition processors, shears and demolition grapples. Before entering into any kind of demolition, number of components is at sight to remove and reuse most of them in the same building during its construction or for the maintenance purpose. Such materials having leftover depreciated value may include sanitary items like taps, copper pipes, GI pipes, conduits for taking electrical wires including all kinds of electric wires, telephone cables and components like electric lamps, bells, attachments for all kinds of electric gadgets etc.. All such materials can be safely extracted using simple machinery and tool kits, small shears and grapples on skid-steer loaders and excavators. A mini-excavator can work at any suitable angles for soft extraction of the items on a vertical surface, while for horizontal surface, a skid-steer can used for breaking up the harder surface. A skid steer equipped with a bucket is suitable for collecting up the useful materials and pushing the same out of the building under demolition.

4.1. Hydraulic Breakers

Hydraulic breakers are available in almost every size as per the need and type of demolition required. Such machinery is suitable for areas which other machineries or tools are not easily and safely accessible to carry out the demolition job. These machineries are suitable for breaking hard surface like RCC slabs and floorings

4.2. Hydraulic Concrete Splitters



Fig-3. Hydraulic Concrete Splinters used for splitting the hard rock

These machineries are suitable for carry out the demolition RCC walls, concrete roads, deck slabs of bridges, industrial floorings or even splitting the rocks into pieces. For using these machines, initially holes are drilled into the surface at certain distances and the tool is pushed mechanically into the holes which split the structural components into smaller pieces.

4.3. Hydraulic Concrete Crushers



Fig-4. Jaw of Hydraulic Concrete Crusher (left) and VHC -30 in operation (Right)

Hydraulic concrete crusher uses the jaws to crush the structural components to be demolished. The jaw opens and closes under hydraulic control and this action helps in crushing the bigger sized components crushed into pieces. The pair of jaws is detachable and suitable types of jaws like cracking jaws, shear jaws or pulverizing jaws are attached as per the requirement. These crushers are mainly deployed as supplementary measures for finer crushing after carrying out the initial demolition by hammers, crushers, blasting, ball and crane, or sawing.

4.4. Long Reach Excavator

The system has a long reach boom which can be operated at non accessible locations at heights up to 20 m and horizontal coverage of 30 m. such machineries are used to pull down the high rise structure in a controlled fashion unlike the wrecking ball.



Fig-5. Komatsu PC450 with super long reach boom and drilling rig attachment.



Fig-6. Caterpillar CAT 385 C

4.5. Diamond Wire Saw

Diamond wire saws are more efficient than circular saws and can cut concrete of any thickness. The wire is wrapped around the concrete component to be cut into pieces with the help of diamond tools.



Fig-7. Diamond Wire Saw with Diamond Tools

4.6. Pulverizers



Fig-8. Pulverizers being used for removing boulders

Pulverizes or grapplers are widely used for removing large boulders and rubbles.

4.7. Rotary Drum Cutters

Rotary drum cutters are very versatile and are widely used for trenching hard stone, cutting pile heads and are put into use for construction as well as demolitions works as for cutting it can cut through concrete as well as bituminous road surfaces.



Fig-9. Rotary Drum Cutters in operation cutting concrete Roads

4.8. Expansive Demolition Agents



Fig-10. Expansive Demolition Agent being poured in predrilled holes (left) and crack developed and widened thereafter (Right)

The expansive demolition agent is a cementitious powder which is mixed in a bucket and poured into the pre-drilled holes. The mix expands after getting hardened leading to the development of hairline cracks in the concrete slab. These cracks widen further and breaking the slab turned into pieces. Biggest advantage of such agents is the elimination of labor component

from the job during the demolition process as no labor is further required after pouring the mix into the holes till the rubble is to be transported away. Other advantages are that there is no noise, no vibrations and no dust is produces in its action.

4.9. Pneumatic Hammer

Pneumatic hammers are mounted on lighter carriers and with the external air compressor are very useful in confined areas for drilling or chipping tough materials such as concrete and asphalt surfaces



Fig-11. Pneumatic Hammer mounted on carrier (left) and being used for drilling rock (Right)

4.10. Hand Held Concrete Cutter

Very handy equipment but heavy duty designed to quickly break away large, thick pieces of concrete. The benefits are limited dust, support a clear view at the rescue scene, no vibrations and prevent risk of secondary collapse.



Fig-12. Concrete Cutter in Hand cutting concrete

4.11. Modular Pile Cutters



Fig-13. Modular Pile Cutter (left) and cutting Pile Head (Right)

These versatile modular pile cutters enable you to hire the exact pile cutter to suit square or circular piles in the size and diameter of your choice. For square piles 200 – 500 mm and circular piles 200 – 900mm. Cutting piles and preparing ready for foundation work has never been more efficient.

5. REUSE AND RECYCLING POTENTIAL OF DIFFERENT C&D WASTE PRODUCTS (Noguchi, 2012; Bansal and Singh, 2014)

The amount of C&D wastes in India has been estimated to be 15 – 20 million T annually with the figure to go further high with the rapid development in infrastructure and exponential growth in real estate. Proportion of concrete is estimated as 23 to 35% of total waste. Considering 30 % percent of C&D wastes as concrete, and 40 % of the concrete as coarse aggregate, the total recycled concrete aggregate (RCA) per year in India can exceed 2 million T.

5.1. Concrete (Radonjanin *et al.*, 2012)

Concrete is primarily a composition of cement, coarse aggregates, fine aggregates and water, further processed by addition of industrial products/ by products for enhancing the strength. Engineers are mainly dependent on nature for obtaining the Coarse and Fine aggregates as well as water for the chemical reaction with cement. Scarcity is there for all these naturally occurring materials and need is there to explore alternative sources. Even for the water with required properties, shift is towards the use of waste water after due treatment. One of the alternative sources of coarse aggregates is recycled concrete aggregates (RCA) which are obtained from the processed Construction and Demolition (C&D) waste. During and after the demolition of any concrete structure, the demolished concrete waste is taken to a recycling plant and there crushed into the required sizes which is called the Recycled concrete aggregate (RCA).



Fig-14. Recycled Coarse Aggregates (RCA) after processing

Sometimes, good sized precast element are also obtained during the demolition, which have a potential of being reused or otherwise, these are also crushed and converted into the recycled aggregates. Limitation is there in use of these recycled aggregates. The production of concrete in India is governed by BIS and IRC codes i.e. IS: 456, IS: 1343 or IRC: 112. All these codes allow the use of naturally occurring aggregates only conforming to IS: 383. To overcome these limitations, it is required to make a special provision of use of Recycled aggregates in combination with naturally occurring aggregates. Thus, use of recycled aggregates can be there with different value of their share by suitable replacing the component of naturally occurring aggregates. It will help out not only in meeting the situation where there is acute shortage of natural resources, but also a steep towards the sustainability. In our bye laws or environmental law, framing of rules will be required for the use of C&D waste. It can have a provision in Environment Protection act, 1986 or a separate act can be introduced independently. Guidelines are required to be framed for use of C&D waste. If we look at other countries like Norway, Japan and Korea, a major junk of demolition waste is recycled and is being used as a partial replacement of natural aggregates and concrete thus produced is being widely used in these countries.

5.2. Brick

Bricks are the important building material in the construction of residential as well as non – residential buildings. It is also a significant component of the total C&D waste on new residential construction sites. Its demand figures next to concrete as a building material. Bricks are largely treated as waste when broken or damaged from the brick production line or from construction site due to poor internal handling and excessive cutting (see Fig. 2). Brick is a maintenance-free component of the structure which is durable during the complete service life of the building. The high durability property of the brick makes it environmental friendly in the sense that after the demolition of the structure, it can be reused repeatedly and the left over volume which is non-reusable can be recycled for other beneficial purposes. Generally, a building is not required to be demolished due to deterioration in the Engineering properties of the bricks. It is for different reason or different needs other reasons that the building has outlived its useful/economical service

life and required to be replaced with new structure. During the demolition process itself, bricks obtained are stacked for next use in its original form after the removal of mortar which is chiseled out and make the brick ready for reuse or recycling, if not reusable.



Fig-15. Demolished Brick

Bricks, after the removal of the stuck up mortar remain reusable for restoration or for new homes and projects. Recovered bricks can be used like a fresh lot of bricks without any further processing. These can also be laid on as brick pavers or for landscaping or any other artistic creations. Brick paved streets are aesthetically pleasing and rain water also percolates through the pavement. Also, a brick surface is cooler in hot months. These street advantages make bricks a good choice in driveways. Bricks on edge are also sometimes used as economical pavement solutions in smaller compounds. Bricks which cannot be reused directly can be disintegrated into smaller sized aggregates or brick chips to be used as construction materials. These recycled bricks products are strong and durable enough in comparison to the original brick. Bricks from demolition sites can be recycled as Road base and construction fill and also as light weight concrete.

Construction debris consisting of bricks can be recycled into brick aggregate through screening, crushing, re-screening and blending, which can then be used as pavement base material by proper mix proportions with cement and fly ash. Brick waste which are not suitable for recycling into the pavement base materials can be used in construction/land fill. Concrete prepared from crushed brick aggregate has good engineering and also better thermal properties but has greater shrinkage than ordinary concrete. Sometimes, during the manufacturing of bricks, due inadequate burning, or sometimes due to over burning, whole lot is turned into the production waste. Though this waste is different from construction and demolition waste, but it can also be recycled like C& D waste and can be suitable used for production of precast elements like paver blocks, curb stones, interlocking tiles by mixing with cement and using as a concrete mix.

5.3. Tiles

Generally, it is difficult to extract tiles from the walls in proper shape and size in order to find them suitable for reuse. It also depends upon the type of the tiles, their life span and the existing conditions. Seepage behind the walls due to leaking water pipes makes them totally non usable. Still tiles extracted from walls, even if these are broken pieces, provide an excellent opportunity to

the artists /designers for making murals or other decorative master pieces. Broken tiles can also be used aggregate after crushing.

If the tiles can be extracted or removed from the wall in good shape and size, these are reused for the same purpose after the removal stuck up mortar and then glued with suitable adhesives available in market today. Creative items like artifacts, table tops, special effects in drive ways, pedestrian subways etc. can be smartly created by reusing for a wide variety of projects. Nek Chand's Rock Garden in Chandigarh, which is internationally renowned, is a perfect example of such reuses. The broken tiles can be further crushed into smaller sizes and can be a partial replacement of gravel and crushed stone in making concrete.

5.4. Timber

The waste timber is not only produced from the demolition of the building, but also from construction of wooden building wherein lot of timber waste is generated. Each source has its own system of recycling and reuse of recovered timber from the demolition of a building or the construction of a building. Whenever a building is decided to be dismantled, timber products like doors and windows are the items which are removed as first step and that too in original form. Timber products have a quality of a long service life which is much longer than the life of the building itself. Hence, in general such products unless eaten by the termites or damaged due to fire do not lose the Engineering properties for a long time and can be used multiple times and thus an environmentally friendly product.

Sometimes, the waste timber cannot be recovered in its original form or is not usable in same shape and size. Such timber can be recycled into new particleboard, medium density fiber boards, animal bedding or can also be used to make renewable energy. Timber used for recycling has to be free from any other demolition products like concrete, mortar, aggregates, sand, bricks, plastic, metals, tiles etc. Wood chips are produced from good quality wooden waste such as large size lumbers. Some of the particleboard producing companies and the pulp and paper producing companies are still using the recycled chip for their products. Chipped or shredded wood is also used as a sewage sludge bulking medium and other products lime pallets.

5.5. Metals

Amongst the metals, steel and Aluminum are the two major products obtained as waste during the construction as well during the demolition of a building. Structural steel obtained during the demolition of a steel structure or left over steel during the construction can be reused directly without much processing. The members can be resized as op the requirement and can be reused directly. Aluminum scrap can be put into reuse by the solid bonding process. If a care is taken in initial stages i.e. during designing with a valid deconstruction plan, then the reusable scrape can be increased to a much greater extent like house hold appliances, without taking the routing the scrap through a foundry. Reusing a steel beam its existing form is better than re-melting it and rolling a new steel beam, i.e. the energy used to re-melt the beam is saved.

Steel waste occurs during the construction and refurbishment of buildings and when they are ultimately demolished and the material becomes available for recycling. Waste from the

manufacture of steel construction products can be easily collected and segregated for recycling. Steel generates almost nil wastage on the construction site. Waste steel which is reusable is equally good in durability criteria and the quality is also well maintained while making products like fire hydrants, steel furniture and also ecologically sustainable.

As far as Aluminum is concerned, it is recyclable multiple times and is always on demand with the need to preserve the environment. Our raw materials vary considerably based on whether we are using primary or recycled aluminum. We must take into account the different sustainability impacts of sourcing primary aluminum from the mined substance bauxite, or recycled aluminum from either pre- or post-consumer sources. Recycling scrap aluminum requires very less energy in comparison to the energy requirement of new aluminum. Because aluminum is infinitely recyclable, it can be reused in applications vastly different from its previous purpose, and it can also be recast into its original form. These properties make aluminum an ideal material for use in premium applications, even after being recycled many times.

5.6. Plastic

Scrap or waste plastic recovered from demolition or construction site is reprocessed and transformed into the entirely different useful products. Typically a plastic is not recycled into the same type of plastic, and products made from recycled plastics are often not recyclable. When compared to other materials like glass and metals, plastic polymers require greater processing to be recycled. The most-often recycled plastic HDPE (high-density polyethylene) is reduced to plastic lumber, tables, roadside curbs, benches, truck cargo liners, stationery (e.g. rulers) and other durable plastic products and is usually in demand. Other application of recycled plastic is in the preparation of a road surface that includes recycled plastic: aggregate, bitumen (asphalt) with plastic that has been shredded and melted at a temperature below 220° C (428 °F) to avoid pollution. Such road surfaces are very durable and monsoon rain resistant.

5.7. Excavated Material

Excavated materials are many times contaminated and require special handling and disposal. It may include hazardous as well as non-hazardous material. Excavated contaminated material that can be re-used will be decontaminated prior to re-use, or if not suitable for re-use will be transported to appropriate treatment facilities or approved landfill sites. If the excavated materials are re-used as road fill base material, then cut and fill quantities should be balanced. It will result in eliminating the off-site disposal. Excavated soils will be retained on site for re-use as backfill while hard rubble will be crushed and re-used on site. Unsuitable material for engineering fill can be used for landscaping.

5.8. Asbestos

Generally, asbestos is disposed as hazardous waste in landfill sites. The demolition of buildings containing large amounts of asbestos based materials have to be deconstructed piece by piece or the asbestos has to be removed carefully before the structure can be demolished. Asbestos can be recycled by transforming it into harmless silicate glass, porcelain stoneware tiles, porous

single-fired wall tiles, and ceramic bricks. Current removal procedures require a completely sealed area, using vacuum to prevent any particles from escaping. Workers must wear heavy protective equipment. It is important to ensure that asbestos waste has been wetted and sealed in heavy-duty plastic prior to transportation to an approved landfill.

5.9. Asphalt Concrete

Demolished asphalt concrete can be utilized as aggregates for asphalt concrete. Also, the demolished asphalt concretes can be used for land fill.

6. SIMPLE GUIDELINES FOR RECYCLING OF DEMOLITION WASTE (Noguchi, 2012; Bansal and Singh, 2014)

The agencies responsible for generation of wastes should separate the generated wastes having potential for reuse/recycling. The Engineer-in-charge will select structure's type and materials that are suitable for reuse/recycling, use recycled aggregates, and ensure proper treatment of wastes generated from such development. The waste generation from construction should not only be minimized, but should also minimize the hazardous effect from the generated wastes.

6.1. Various Agencies (or Sub-Contractors)

Various agencies or sub-contractors to be involved are to be linked up with the steps in this process of C& D waste reuse and recycling. Some of such steps can be listed as waste collection and transportation, waste middle treatment i.e. receiving the waste, its segregation and further suitable comprehensive treatment before putting into the use. The cost for C&D waste separation, storage, treatment, reuse/recycling should be included in the Estimated Cost by the Engineer-in-charge while according Technical sanction and preparing tender documents.

There are important duties to be either assigned or as a dutiful contractor, he may be establishing himself like, he should establish step-by-step demolition plan. Contractor may establish treatment facility at site only. He should report expected amount of wastes by type and treatment plan at the beginning of construction. There should be effective utilization of recycled aggregates and Safe treatment of hazardous waste like asbestos.

6.2. C&D Waste Information on Web

All C&D waste information by contractor and by those involved in its treatment waste treatment companies are to be put on public domain in order to improve the rate of use of demolished concrete for e.g. application of recycled aggregates. Further to have a stronger data base of C&D waste, users reusing the C&D waste or recycled waste after treatment and processing can contribute a lot. This will help in substantial reduction on the amount of wastes and promotion of recycling or reusing the C&D waste.

6.3. Demolition Plan

It is required to adopt a systematic approach while demolishing a building in order to minimize the waste and its best use. A recommended approach can be to follow a sequence of segregation of

household waste as first step followed by mechanical and electrical equipment, exterior and interior finishing materials, roof finishing and water-proofing materials, then structure as a last resort. Demolished C&D wastes need to be brought out of field immediately or temporarily stored in a designated area for the C&D wastes.

7. RECYCLING THE DEMOLISHED STRUCTURE

Once the structure is demolition as per the planned and in an Engineering manner using the most appropriate tools and machineries, the next step is the segregation of the products which can be put into reuse rather than recycling to useful aggregates or dumping in safe and approved manner.

Initially the useful products like bricks, steel frames, aluminum windows, wooden frames, doors and windows having reusing potential are segregated. There may be some material which are having good scrap value are also segregates before the debris is sent to the recycling plant. Even the reinforcement bars are taken out of RCC and may be put into the use in non-sensitive places like kitchen slabs, lintel beams.

For a proper and an efficient functioning of products in a recycling plant, it is essential that the debris brought to the plant is free from contamination by other types of undesired products like plastic, oils and paints. Use of appropriate equipment and machinery is essential in recycling process. Such equipment may be jaw crushers, magnetic separators, vibratory screens, washing equipment and can be deployed to segregate such scraps. Few of such machines are discussed in the following paras.

7.1. Portable Concrete Crushers

It crushes concrete into coarse / fine aggregate, which can then be recycled onsite. After pulling down the structure, the concrete slab is cut into smaller pieces to run through the crusher. The machine then grinds it into gravel-sized aggregates, which can be used for road compaction, structural fill, pipe bedding and other applications.



Fig-16. Portable concrete crushers in operation (left) and on mounted on carrier (right)

7.2. Track Mounted Mobile Recycling Crusher

Mobile crushing machine is usually utilized for contracting, quarrying and crushing work under mine. The mobile crusher can crush any types of rocks and minerals. The hydraulic driving track mounted mobile crushing machine are extensively utilized in road and bridge construction, urban construction, metallurgy, energy as well as other departments to do crushing and sieving perform.

7.3. Jaw Crushers

Jaw Crusher (or Rock crusher) is ideally suited for primary crushing of stones and is often used in mining, metallurgy, building materials, roads, railways.



Fig-17. Jaw Crusher in operation (Right)

7.4. Mobile Jaw Crushers

Mobile jaw crushers are suited for small to medium sized and have hydraulic adjustments for changing the size of finished products. The primary crushing applications ranges from hard and abrasive to mixed recycled materials.



Fig-18. Moblie Jaw Crusher

7.5. Cone Crushers

A cone crusher is generally used as a secondary crusher in a crushing circuit. Pre-crushed



Fig-19. Cone Crushers

Product (usually 3in minus depending on the cone crusher model) is fed through the top of the cone crusher and flows over the mantle. The vertical cone crusher drive shaft rotates the mantle eccentrically below the concave or bowl liner, squeezing the product and crushing it between the mantle and concave. Cone crushers are usually run on belt drives driven by an electric motor or diesel engine. Cone crushers are used extensively throughout the aggregate and mineral processing industry.

7.6. Mobile Cone Crushing Plant

The mobile cone crushing plant works on construction site. The mobile cone crushing plant can be shifted to crushing place directly as it will maneuver both in smooth and bumpy roads. It is compact to use and operations are simple. The mobile crushing plant fitted with generator, motor and control box on the trailer is suitable to work at outdoor with ease. The mobile crushing unit can function independently or can be used as a secondary unit for fine crushing. It can also work as a production line with the screening equipment according to the requirement of customers. Materials are delivered by belt conveyors, which is simple to operate and easy to maintain.



Fig-20. Mobile Cone Crushing Plant in operation

7.7. Horizontal Shaft Impactors

Horizontal Shaft Impactor (HSI) crushers break rock by impacting the rock with hammers that are fixed upon the outer edge of a spinning rotor. These machines are available in stationary, trailer mounted and crawler mounted configurations. They are used in recycling, hard rock and soft

materials in wet and dry applications. Earlier it was used on soft materials and non abrasive materials, such as limestone, gypsum etc. They are used as primary and secondary crushers for size reduction. These Crushers offer operators a very high ratio of reduction and can accept larger feed sizes.



Fig-21. Horizontal Shaft Impactors in operation of breaking rock (Right)

7.8. Screens

Screens are equipment that is used for the separation of material into 2 - 6 different sized products. The material is separated by passing it through a vibrating 'screen box' which a number of different sized screens, or meshes, which the material falls through like a sieve.



Fig-22. Screening of materials

The material falls onto attached conveyors that stock the end products which can then be used in the building and construction industries. Screening equipment is either track or wheel mounted to make them mobile. The equipment can move to the raw material rather than the expensive movement of the material to the screen.

7.9. Conveyors

Conveyors are designed to work in conjunction with the entire range of washing and screening equipment. Conveyors can be electrically or hydraulically driven. With the mobile range the ability to radial increases stockpile capacity and reduces on-site material handling.



Fig-23. Conveyor

7.10. De Watering Screens

Dewatering Screens are used for de-watering, de-sliming, de-gritting, rinsing, scrubbing and washing. Dewatering Screens offer multiple advantages for construction and specialty aggregates producers including a drier, drip-free product that other types of equipment cannot provide. There are many applications for Dewatering Screens in the aggregates market including dewatering after Sand Classifying Tanks and Fine Material Screw Washers, or use within a system such as the Recipe Sand Plant.



Fig-24. Dewatering Screens

7.11. Bucket wheel Screens

This mobile screen recovery screen has twin bucket wheel units for the removal of clays, silts and slimes to produce 2 grades of sand. It can support maximum retention or disposal of fines. The key advantage of the machines is their ability to dewater sand, with maximum removal of water and at the same time retain maximum amount of fine saleable product.

7.12. Washing Plant



Fig-25. Washing Plant in operation (left) and Power screen (Right)

With an aim to remove clay and silt from sand and gravel, washing plants are required to be deployed at all recycling plants. Power screen T4026 helps in washing the sand and gravels to produce cleaner aggregates and has a special feature of recirculation of the water

7.13. On-Site Bucket Crushers

On site Bucket Crushers can be put on job directly at site for processing and recycling of concrete aggregates and bituminous overlays. No separate crushing and handling is required which makes it cost effective by saving the cost of transportation also and easily installable on-site.

7.14. Materials Sorting Screens & Steel Separation Magnets

Typically, mixed C&D debris is tipped at a central facility, and the materials with a high market value, such as large pieces of sawn lumber, are removed. The remaining mixed C&D materials are then processed using one of two primary methods. The mechanized size reduction method uses a crusher, a dozer, or a compactor.

The materials are then passed through a series of screens, magnets, and other separation equipment. The manual labor method relies on human sorters to pick out materials and place them in specific containers.



Fig-26. Material Sorting Screen

Screens and magnets may also be employed with the human labor method, but the materials are left in their original form rather than crushed so that they can be easily distinguished and sorted. The most common approach is a blend of the mechanized size reduction and the human sorter methods.

7.15. Concrete Recycling Plants

Specialized equipment is also available to reclaim slurry water and aggregates for reuse in concrete production these days. The demand for such eco-friendly equipment is on the rise. Slurry water is often used in concrete batching plants. With the massive infrastructure development activities going on around the country, one can see increased sightings of concrete recycling plants these days.

While concrete along with other aggregates is recycled, the same holds good for steel and other ferrous scrap that is generated from demolition projects. The salvaged material finds its way to steel mills where it is put through a series of processes in order to make them ready for reuse in construction projects.

The key to the success of recycling lies in proper planning with a comprehensive waste management plan. Developers need to formulate a plan to first of all segregate the various types of waste that is generated from demolition projects. This ensures a smooth work flow and sustainable use of demolition waste.

By building implosion or controlled demolition method, a structure can be demolished in seconds, which otherwise is cumbersome and time consuming. This method is widely used in urban areas for demolishing large structures in the developed countries. Many structures in India were also demolished using implosion techniques. Structures like chimneys can be demolished in a short span of time sometimes even less than a day. Larger structures may take up more time in preparation such as to remove columns and walls before firing the explosives.



Fig-27. Concrete Recycling Plant

The controlled demolition is done by placing a series of small explosives in strategic locations in the structure and its detonation is so timed that a structure collapses down with the gravity on itself, minimizing any damage to its surroundings. Most commonly used explosives are Nitroglycerin, dynamite to blast reinforced concrete supports in the structure. Explosives are then detonated progressively throughout the structure. This method is also used for demolition of other structures as well other than buildings.

8. C&D WASTE RECYCLING FACILITY IN NEW DELHI (IL& FS ECOSMART)

About 5000 T of C& D waste is generated in Delhi every day. Out of which 90 % is due to Re-construction activities requiring demolition of existing structures and just 10% is contributed due to new constructions taking place. Out of the total C&D waste, hardly 50 % goes to the landfills, 10 % is recycled and balance 40 % is getting illegally dumped causing all kinds of nuisance, polluting the environment, natural piece of lands and water bodies.



Fig-28. Illegal Dumping of Demolition Waste at various Location in Delhi

Types of waste received at landfills include soil/ sand, gravels, bitumen, metals, bricks, concrete wood, plastics. Most of the metals, timber, glass and plastics is almost collected separately

and recycled. Out of all the constituents of total wastes, masonry (bricks as well as concrete) items and soil/sand are the main constitutes to the order of 90 %.



Fig-29. C&D Waste Management Plant setup by IL&FS at Burari, Delhi.

One pilot plant set up IL&FS Environmental Infrastructure and Facilities Ltd. for North Delhi Municipal Corporation is functioning at Burari, Delhi. The plant is built up in about 7 acres (28000 sqm) of land with a capacity to handle about 500 T of C&D waste per day.

8.1. Collection and Transportation of C&D Waste

There are designated collection points where containers and skips are placed as a first storage point of C&D waste where the waste is brought in by public and private persons. As an additional measure, sufficient vehicles re also in operation to collect the waste from various locations of the city as per the services required from pick up points. After the collection of Construction and Demolition waste, the material is routed through the Weighbridge to the processing site.



Fig-30. Skips and Containers

8.2. Processing Site

First of all silt, loose soil and muck is separated in the processing site by screening through a grizzly. This screened loose soil or muck makes more than 50% of the waste and the only option is to utilize it for landfill. In order to make an effective use of this leftover portion, pure sand is extracted by processing it wet processing system called Evowash system.



Fig-31. Evowash System

Next stage of screening is sorting of bricks and concrete manually. Before actual processing starts, the bigger sized masonry or concrete wastes are reduced to smaller size in a rock breaker. Then the leftover waste is processed in central processing unit which has mobile crushing units Rubble Master RM 60, with capacity of 60 T per hour machines for further reduction in size. This Processed C&D waste is used for sub base of roads and for making bricks, paver blocks & kerb stones. The process flow of C&D waste recycling at the plant is as under:-

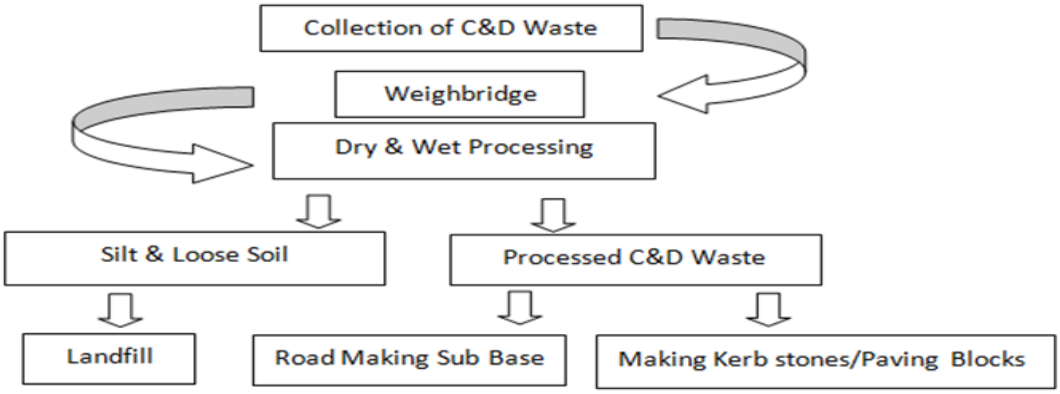


Fig-32. Flow- Diagram of process of C&D waste recycling

8.3. C&D Waste Utilization Test Road



Fig-33. Test Road with C&D Waste

8.4. Pavement Blocks and Kerb Stones



Fig-34. Pavement Blocks (Top Left), Bricks (Top Right), Tiles (Bottom Left) and Kerb Stones (Bottom Right) made from C&D waste

9. CONCLUSION

Concrete has become a high tech material and its production is used by economists as a measure of a country's economic strength. Aggregates from natural sources are not available. The situation has forced us to explore aggregate from alternate sources. It is important to minimize C&D waste generation and maximize reuse/recycling as the construction industry is consumer of tremendous amount of natural resources.

In so far as use in concrete in India is concerned, IS: 456, IS: 1343 or IRC: 112 do not permit use of aggregate other than those obtained from natural sources and conforming to IS: 383. In view of international developments and experiences, as well as shortage of aggregate from natural sources being experienced in many parts of the country, it is time that recycled aggregate are permitted for use in concrete constructions.

The Municipal bye-laws are required to be reviewed and suitably modified. Establishment of effective strategies and enactment of laws and regulations deem to be essential to achieve this. In addition, provisions of some incentives to users of the recycled products deem to be necessary to promote the use.

Research & Development is to be promoted by Government. There is lack of public awareness and it is required to spread the Information, Education and Communication in order to Garner Public Support and change the attitude of Public and Staff. Data and results should to be posted in Public Domain.

10. ACKNOWLEDGEMENT

The authors are highly thankful to Indian Concrete Institute, New Delhi Centre for organizing a workshop on Construction and Demolition (C&D) waste recycling on February 28 and March 01, 2013 in New Delhi. The three international speakers namely Prof. Takafumi Noguchi from Japan, Dr. Dong uk CHOI from Korea and Dr. Christian J. Engelsen, Senior Scientist, SINTEF from Norway have shared their international experience with Indian delegates. Further authors are thankful to Dr. A K Mullick, Chairman of C&D waste recycling committee of ICI for providing valuable information on the subject. Further authors are thankful to Dr. N B Majumder from

IL&FS and Er. Pradeep Khandelwal, Chief Engineer, EDMC for providing vital information regarding the pilot plant setup at Burari for processing of C&D waste.

REFERENCES

Bansal, S. and S.K. Singh, 2014. A sustainable approach towards the construction and demolition waste. International Journal of Innovative Research in Science, Engineering and Technology, (An ISO 3297: 2007 Certified Organization), 3(2): 9226-9235.

Dhir, R.K. and K.A. Paine, 2010. Value added sustainable use of recycled and secondary aggregates in concrete. Indian Concrete Journal, 84(3): 7–26.

IL& FS ECOSMART. Construction & demolition (C&D) waste; collection, transportation and disposal system. Project Report for MCD, Delhi Solid Waste Management Program: pp: 38.

Mullick, A.K., 2012. Green options for binder system and aggregate in sustainable concrete. Indian Concrete Journal: 9–17.

Noguchi, T., 2012. Sustainable recycling of concrete structures. Indian Concrete Institute (ICI) Journal, April – June, 13(1): 40–53.

Radonjanin, V., M. Malesev and S.A. Marinkovi, 2012. Recycled concrete as aggregate for structural concrete production. The Masterbuilder, Chennai: 58–72.

BIBLIOGRAPHY

Bansal, S. and S.K. Singh, 2013. Recycling and reuse of construction & demolition (C&D) waste: Technology and equipments. 6th Engineering Conference in the Asian Region and Annual HAKI Conference, Embracing the Future through Sustainability 20-22 August 2013, Jakarta. pp: 102.

Bansal, S., S.K. Singh and J. Kurian, 2014. Construction & demolition (C&D) waste recycling in New Delhi. 4th International FIB Congress 2014. Mumbai, India February 10 to 14, 2014, Improving Performance of Concrete Structures, Proceedings (Summary Papers Vol. 1). pp: 286-289.

ELK
Asia Pacific Journals

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

TENSILE BEHAVIOUR OF 3-PLY LAMINATE COMPOSITE OF SHEET METALS

Vijay Gautam, Assistant Professor, Department of Mechanical Engineering, Delhi Technological University, Delhi-110042. Email: vijay.dce@gmail.com	**Bijender Prasad**, Research scholar, Department of Mechanical Engineering, Delhi Technological University, Delhi-110042. Email: bijendra121@gmail.com

Abstract- **Recently, clad metallic materials, consisting of two or more layers, have been preferred in various industrial applications because of their unique corrosion resistance, specific strength, and surface properties. The present study has been carried out on three ply composite laminate, containing AISI304L austenitic stainless steel on one side and AISI430 on the other side with AA1050 in the core of the blank. Apart from excellent corrosion and mechanical properties of stainless steel three ply composite laminate possesses exceptional thermal and electrical conductivities, which makes it useful for utensils formed by deep drawing process. Circular blanks of composite laminate, produced by cold roll bonding, with a combined thickness of 2.5mm, were procured in the annealed condition from a leading manufacturer. To ensure the bond strength of 3-ply AISI304/AA1050/AISI430 sheets, peel tests on various specimens were performed according to ASTM-D1876-08 standard. Tensile samples as per ASTM E8M standard, were laser cut in three different directions i.e. parallel, inclined at 45° and transverse with respect to rolling direction. Tensile specimens were tested to study the deformational behaviour in uniaxial tension, on a 50kN UTM.**

Keywords- ***Clad metals; Three ply laminate composite; cold roll bonding; peel test; tensile behavior.***

I. INTRODUCTION

With the advancement in technology in forming with sheet metals, new clad materials have been evolved and designed for various industrial applications such as automobile, aerospace and electrical industries. These laminated sheets consist of different kind of sheet metals with different mechanical, physical properties and specifications to suit the various applications. Various parameters such as material type, thickness ratio, arrangement and multiplicity of the metals, surface preparation, bonding parameters and post heat treatment results in unique deformational behavior [1-5]. Although, Clad sheets have been produced by several solid state bonding methods such as diffusion, explosive and roll bonding but cold roll bonding (CRB) is the most efficient and cost effective [6,7]. Many researchers have contributed significantly on the various issues related to formability of roll bonded clad sheets. In the recent development of the clad sheets, most widely used material combinations are Al with Cu, Al-Fe, Al with stainless steel, Zn- Al with stainless steel, Al-Cu with steel, Ti with steel etc. [8,9]. Akramifard et al 2014 carried out experimental studies on effect of reduction and subsequent annealing temperatures on mechanical properties and bond strength of three layered AA1050-304L-AA1050 clad sheets, during roll bonding. An important contribution of their work was the correlation of tensile and peel test on the basis of principles of mechanics of materials. The mechanical properties of some laminated composite sheets, mainly of stainless steel/aluminium sandwich sheets, have been the subject of examination for many years. Choi et al. 1997 experimentally investigated the deformation behaviour of Al-STS430 bi-layer clad sheets under uniaxial tension and concluded that a difference in planar and normal anisotropy results in the warping of edges of tensile specimens. The material property of the laminated sheet changes in the thickness direction, their deformation behavior should be affected by the blank placement position during forming by deep drawing, i.e. which side of the sheet would contact punch or die. Because of such interesting formability issues, the behavior of laminated sheets is the subject of much research. However, only a few studies have discussed the combination of 3-ply clad sheets of AISI304/AA1050/AISI430. In view of the above the present study deals with the tensile behaviour of 3-ply laminate composite.

II. METHODOLOGY

A. Selection of material

Selection of material is based on the increasing popularity of multi ply clad sheets in deep drawn utensils, owing to the advantages of excellent formability, good surface finish of class of 2B, uniform heating and better heat transfer. Circular blanks of 3-ply clad sheet of AISI304/AA1050/AISI430 material of effective thickness of 0.4, 1.5 and 0.6mm respectively, that are

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

joined by roll bonding and thermal treatment, due to a metallurgical bond under high pressure, was procured from a leading supplier in India. The chemical composition of individual sheets were obtained by spectrometry- analysis and are given in Table 1.

B. The peeling test

To investigate the bond strength due to CRB, the peeling test was performed according to ASTM-D1876-08 standard (subsized). Specimens for the peel test were laser cut in the sub-size of 160X 25 mm due to limited availability of the material. Since it is difficult to detach the bonded sheets mechanically, one end of the each specimen was immersed in the solution of sodium hydroxide to dissolve aluminium. The remaining length with intact bond was used for the peel test as shown in the Fig. 1. The specimens were prepared in such a way that each bonded sheet of steel i.e. AISI304 and 430 be peeled from aluminium bond respectively. The stainless steel sheet to be peeled was held in the lower fixed jaw and the rest with the upper movable jaw fixed with the cross -head on the 50kN UTM. All the tests were conducted at a cross head speed of 10mm/min.



Fig. 1. Peel test arrangement on 50kN UTM

C. The tensile test

Most common approach to characterize the behaviour of a material is by conducting uniaxial tensile tests. In this work, simple tension tests were carried out on a universal testing machine of maximum capacity 50 kN as shown in Fig. 2.

The tensile test specimens as per standard ASTM-E8M, were prepared by laser cutting of the blank as shown in Fig. 4. The anisotropy of the 3-ply clad sheet metal was investigated by performing tensile tests at specimen orientation of 0°, 45° and 90° to the rolling direction (RD) and are shown in Fig. 5. The tests were carried out include monotonic loadings in tension. Each test is performed at least three times to ensure good reproducibility of the experiments. The tests were carried out at a cross head speed of 2.5mm/min. Typical engineering stress strain curve is plotted on the basis of force and displacement data acquired from the dedicated software.

D. Determination of tensile properties

The strain hardening exponent (n) and the strength co efficient (K) values are calculated from the stress strain data in uniform elongation region of the stress strain curve.



Fig. 2. 50kN Universal testing machine



Fig. 3. Tensile tested specimens

ELK
Asia Pacific Journals

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

TABLE I CHEMICAL COMPOSITIONS OF SHEET MATERIALS USED IN THE LAMINATE

Steel Sheet material	C	Mn	Si	Al	Ni	Cr	S	P
AISI304	0.05	1.450	0.820	0.044	9.620	18.03	0.009	0.023
AISI430	0.12	0.902	0.675	0.038	0.621	17.02	0.023	0.026
AA1050	0.18	0.018	0.156	Rest	0.032	0.012	0.015	0.034

The plot of ln(true stress) versus ln(true strain) which is a straight line is plotted as discussed below :

The power law of strain hardening is given as :

$$\sigma = K \epsilon^n \quad (1)$$

where σ and Є are the true stress and the true strain respectively.

Taking log on both sides

$$\text{Log}(\sigma) = \log(K) + n \log(\epsilon) \quad (2)$$



Fig. 4. Laser cutting of tensile specimens



Fig. 5. Tensile specimens at different orientations w.r.t. RD

This is an equation of straight line and the slope of which gives the value of 'n' and 'K' can be calculated taking the inverse natural log of the y intercept of the line as shown in Fig. 8.

III. RESULTS AND DISCUSSION

To determine the bond strength of AA1050 with AISI304 and 430 respectively, the peel test specimens were prepared and tested on 50kN UTM of Tenius Olsen make. The peeling force and distance data were recorded using a dedicated software Horizon and the plot between them is shown in Fig. 6.

The average peel strength was determined as the ratio of peeling force to the peeling width. The range of the peeling strength of AISI 304 from AA 1050 was between 30 to 51N/mm, whereas peeling strength of AISI430 from AA1050 was found to be between 20 to 26N/mm. The experimental investigation showed that the peeling strength of AISI304 is higher than AISI430 by a factor of 1.5 approximately. In both the peeling tests, the peeling

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

strength remained stable for an appreciable length of 80mm with a slight dip in between at 40mm of peeling distance.

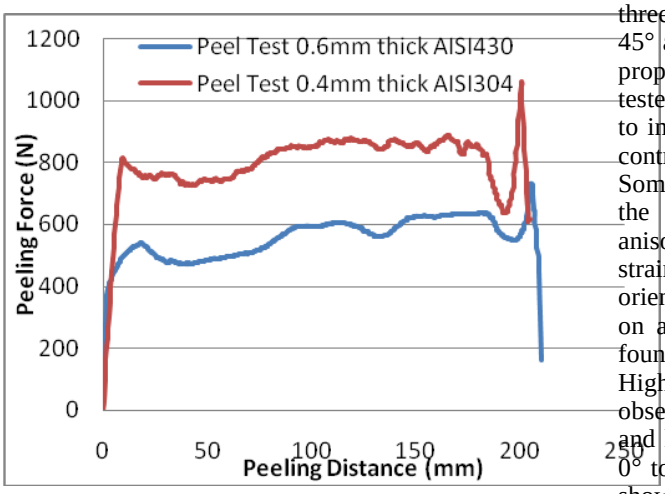


Fig. 6. Curve between Peeling force and Peeling distance

The peeling strength of aluminium from AISI 304 was 40N/mm on an average basis, whereas the peeling strength of AISI430 was 25N/mm. Peeling strength became maximum as the peeling distance reaches near the end of the specimens. Same trend was observed in all the tests conducted.

To investigate the deformational behaviour, tensile tests were performed on the specimens cut at three different orientations i.e. parallel to, inclined at 45° and perpendicular to the RD. The various tensile properties are given in Table 2. In all the tensile tested specimens, nucleation of crack was observed to initiate from 0.6mm thick AISI430 which may be contributed to the lower ductility of this steel grade. Some waviness or warping were seen at the edges of the tested specimens due to the difference in anisotropy of individual sheets. A typical true stress-strain curve is shown in Fig. 7. Tensile specimens oriented at 45° to the RD showed maximum ductility on an average of 57% whereas least ductility was found with the specimens oriented at 0° to the RD. Highest tensile strength of order of 260MPa was observed in the specimens oriented at 90° to the RD and least strength was seen in specimens oriented at 0° to the RD. The strain hardening exponent (n) as shown in Fig. 8, which is an indicator of workability at room temperature was found to be of order of 0.25 on an average basis. High ductility coupled with excellent strain hardening exponent and strength is required in deep drawing of the utensils.

TABLE II TENSILE PROPERTIES OF COMPOSITE LAMINATE AT DIFFERENT ORIENTATIONS wrt RD

Orientation w.r.t. Rolling Direction (RD)	Yield Stress ($0.2\%\sigma_y$) (MPa)	Ultimate Tensile strength (MPa)	Percentage elongation (%)	Strain Hardening Coefficient (n)	Strength Coefficient, (K) (MPa)
0°	171	242	44.7	0.234	420
45°	177	254	56.0	0.256	456
90°	178	262	47.2	0.255	470

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

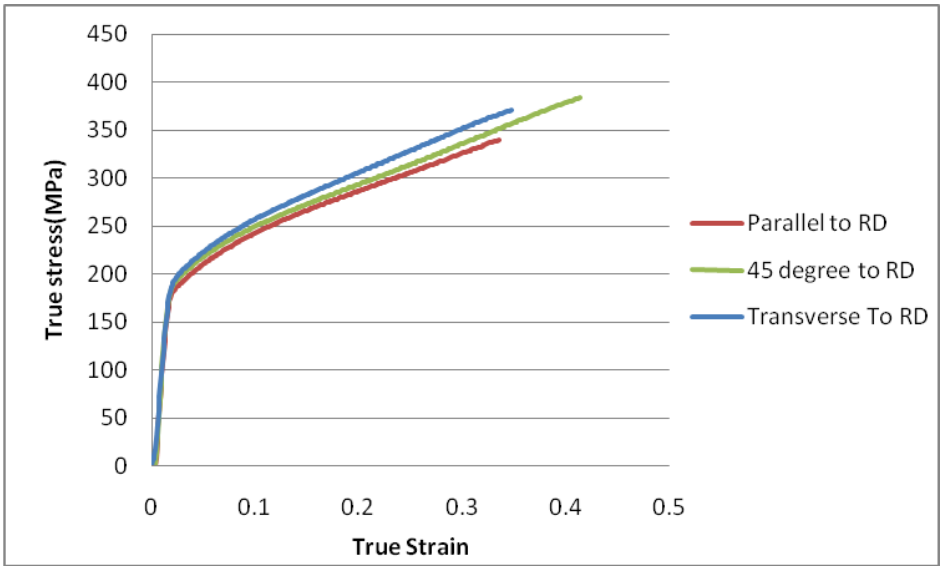


Fig. 7. True stress vs True strain plot for composite laminate oriented at different directions

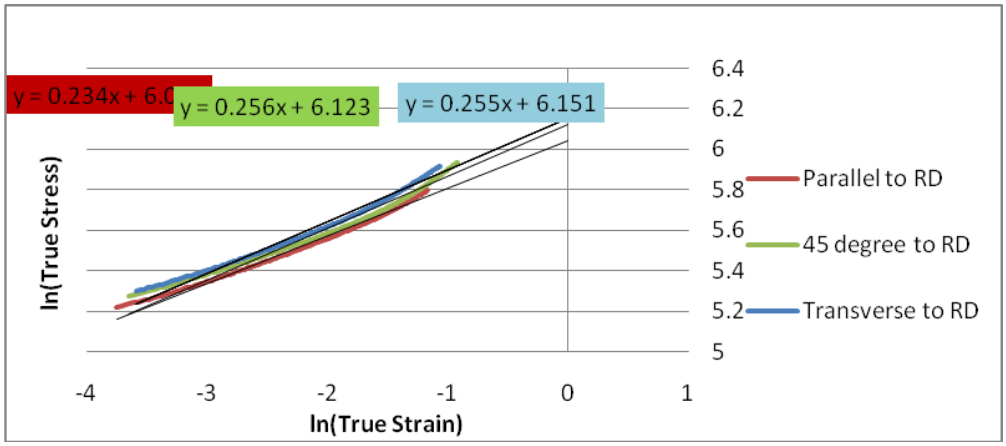


Fig. 8. ln(True stress) vs ln(True strain) plot for composite laminate

IV. CONCLUSIONS

On the basis of the experimental investigations conducted to study the deformational behaviour in uniaxial tension test of 3-ply laminate composite of AISI304/AA1050/AISI430, following conclusions can be drawn:

1. The peeling strength of the bond of AISI 304 with AA1050 is of order of 40N/mm which is approximately 1.5 times higher than that of AISI 430 with AA1050.
2. The peeling strength increases towards the end of the peeling distance.
3. Tensile specimens oriented at 45° to the RD showed maximum ductility on an average of 57% whereas least ductility was found with the specimens oriented at 0° to the RD.
4. In all the tensile tested specimens, nucleation of crack was observed to initiate from 0.6mm thick AISI430 which may be contributed due to the lower ductility of this steel.

ELK Asia Pacific Journals – Special Issue
ISBN: 978-81-930411-4-7

References

[1] J.Y. Jin and S.I. Hong , "Effect of heat treatment on tensile deformation characteristics and properties of Al3003/STS439 clad composite". Mater.Sci.Eng., A 2014; 596:1–8.

[2] L. Chen , Z. Yang , B. Jha , G. Xia and J.W. Stevenson "Clad metals, roll bonding and their applications for SOFC interconnects", J Power Sources 2005;152:40–5.

[3] E.Y. Kim ,J.H. Cho ,H.W. Kim and S.H. Choi . "Evolution of deformation texture in Al/Al– Mg/Al composite sheets during cold-roll cladding", Mater.Sci.Eng, A 2011;530:244–52.

[4] I.K. Kim and S.I. Hong . "Effect of component layer thickness on the bending behaviors of roll-bonded tri-layered Mg/Al/STS clad composites". Mater.Des. 2013;49:935–44.

[5] H.G.Kang , J.K. Kim, M.Y. Huh and O. Engler . "A combined texture and FEM study of strain states during roll-cladding of five-ply stainless steel/aluminum composites". Mater.Sci.Eng, A 2007;452–453:347–58.

[6] I.K. Kim and S.I.Hong. Mater.Des.49(2013)935–944.

[7] H.R.Akramifard , H. Mirzadeh and M.H. Parsa,"Estimating interface bonding strength in clad sheets based on tensile test results" , 2014, Mater.Des.

[8] H.R.Akramifard , H.Mirzadeh and M.H.Parsa "Cladding of aluminum on AISI 304L stainless steel by cold roll bonding: Mechanism, microstructure, and mechanical properties" Mater.Sci.Eng. A613 (2014)232–239.

[9] S.H. Choi ,K.H. Kim, K.H. Oh and D.N. Lee ."Tensile deformation behavior of stainless steel clad aluminium bilayer sheet", Mater.Sci.Eng. A222 (1997) 158-165.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Text Watermarking using Sinusoidal Greyscale Variations of Font based on Alphabet Count

Jeebananda Panda[1] , Nishant Gupta[2] , Parag Saxena[3], Shubham Agrawal[4], Surabhi Jain[5], Asok Bhattacharyya[6]

Associate Professor, Department of ECE, Delhi Technological University, Delhi, India[1]

UG Student, Department of ECE, Delhi Technological University, Delhi, India[2]

UG Student, Department of ECE, Delhi Technological University, Delhi, India[3]

UG Student, Department of ECE, Delhi Technological University, Delhi, India[4]

UG Student, Department of ECE, Delhi Technological University, Delhi, India[5]

Professor, Department of ECE, Delhi Technological University, Delhi, India[6]

ABSTRACT: In this paper, an imperceptible, fragile text watermarking algorithm has been proposed. The font color of each alphabet in the text document is varied in the greyscale in accordance with a sine function. The amplitude for the same is generated using a hash function. The authenticity of the received document can be verified by comparing the actual color values of the letters in the document with the expected values generated by the sine function. This scheme is highly sensitive to various malicious text related tampering attacks hence preserving the integrity of the document. Unlike earlier methods, the proposed algorithm provides tamper detection while maintaining watermark invisibility. The attack analysis illustrates that the scheme is efficient and secure.

KEYWORDS: Greyscale; sine wave; text watermark; digital watermark; hash function; tamper detection

I. INTRODUCTION

The increased role of internet and networking techniques in modern communication has raised concerns over the security of digital information. Greater access to mobile devices like flash drives, memory cards, iPods etc. has enabled large volumes of text material to be transferred over these channels, exposing them to plagiarism, copyright violation, redistribution and other forms of malicious attacks. While extensive work has been done in the fields of image, video and audio watermarking; the research in the field of text watermarking is rather limited. The growth of e-commerce, e-business and digital libraries has augmented the need for efficient text watermarking techniques. Over the years the methods of encryption, steganography and watermarking have been used to solve these problems. Recently digital watermarking has emerged as a more advantageous method. This method is preferable over its counterparts as it maintains the comprehensibility of the documents while ensuring their authenticity and integrity. The proposed algorithm is based on the variation of the colour values of the font which follows a sine function. A hash function is implemented on a paragraph of the text to generate amplitude values for the sine function. The algorithm is sensitive to any form of tampering attack. The paper is organised into 6 sections. Section II examines the previous related work done in the area of text watermarking. Section III describes the proposed algorithm. Section IV illustrates the implementation of the same. Section V describes the experimental results and Section VI lists the conclusions.

II. RELATED WORK

A digital watermark may be described as an identification code that is permanently embedded in the document. The invisible watermarks are more secure at preserving the authenticity of the document. In the past many techniques have been proposed. These include text watermarking using text images, synonym based, pre-supposition based, syntactic

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

tree based, noun-verb based, word and sentence based, acronym based, typo error based methods etc. The text watermarking approaches are classified as follows.

A. *Image Based Approach:*
In image based approach towards text watermarking, the watermark is embedded in the text image. Brassil, et al. were the first to propose a few text watermarking methods utilizing text image [1]-[2]. Later, the performance of these methods were analysed by Maxemchuk, et al. [3]-[4]. Huang and Yan [5] proposed an algorithm based on an average inter-word distance in each line. Wiyada Yawai and Nualsawat Hiransakolwong showed how to use the intersection position of horizontal line, virtually run across text character skeleton line on a text image under the cross ratio applying, to be the marking point of zero watermarks [15].

B. *Syntactic Approach:*
The constituents of a sentence such as nouns, verbs, prepositions etc. determine the syntactic structure of the sentence which depends on language and its conventions. Applying syntactic transformations on text structure to embed watermark has also been one of the approaches towards text watermarking in the past. Mikhail J. Atallah, et al. first proposed the natural language watermarking scheme by using syntactic structure of text [6]-[7]. Hassan et al. performed morpho-syntactic alterations to the text to watermark it [8].

C. *Semantic Approach:*
Semantics of text like synonyms and antonyms are utilized to embed the watermark in text. Atallah et al. were the first to propose the semantic watermarking schemes in the year 2000[9]. Later, the synonym substitution method [10] was proposed. A noun-verb based technique for text watermarking was also proposed [11] which exploit nouns and verbs in a sentence parsed with a grammar parser using semantic networks. Later Mercan, et al. proposed an algorithm of the text watermarking by using typos, acronyms and abbreviation to embed the watermark [12]. Algorithms were developed to watermark the Text using the linguistic semantic phenomena of prepositions [13]. The algorithm based on Text Meaning Representation (TMR) strings has also been proposed [14].

D. *Zero-Watermarking approaches:*
In zero watermarking approach, the host text document is not altered to embed the watermark; rather the characteristics of the text are utilized to generate a watermark. This watermark pattern is later matched using a pattern matching procedure with the pattern generated by tampered document to identify any tampering [16]. Jalil Z. et al. proposed a zero text watermarking algorithm based on occurrence frequency of non-vowel ASCII characters. The embedding algorithm makes use of frequency non-vowel ASCII characters and words to generate a specialized author key [17]. Zunera Jalil et al. developed an algorithm which utilizes a keyword from the text (selected based on author choice) to generate a watermark based on the length of preceding and next word length, to and from the keyword occurrences in text [19].

III. PROPOSED ALGORITHM

The aforementioned techniques have the drawback of attack specificity and often become unreliable when multiple attacks are performed. They also are not applicable to all types of text documents under random tampering attacks and are not specifically designed to solve tamper detection problem. The proposed algorithm aims to ensure authenticity and integrity over a wide variety of tamper attacks while also identification of location of tampering.

The algorithm uses variations in the font colour of the document. This watermarking technique exploits the fact that minor changes in colour are imperceptible to the human eye. The colour of every alphabet in the document varies sinusoidal within the greyscale. Xianmin Wei, earlier, proposed a sine wave based watermarking scheme which relied on word count, was limited to WORD documents only [18].

In the proposed algorithm, the count of each alphabet present in the raw text determines the parameters of the color variations on the entire text. It is, therefore, independent of any format of the document i.e. doc, docx, pdf etc. because the algorithm runs on raw text. The same algorithm may easily be extended to other languages with only slight

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

modifications involving the number of alphabets. The proposed technique of digital text watermarking can be adopted for IP protection of any text document.

A. *Watermark embedding process:*
The amplitudes of the sine waves are generated by applying the SHA-1 hash to any chosen paragraph of the document. This paragraph will be pre-decided between the sender and the receiver. The first 26 hexadecimal bits of the 40-bit hash so generated are used as amplitudes. The first 26 hexadecimal bits are used from the 40 bits generated as the amplitudes. Since these are hexadecimal values, the amplitude corresponding to sine wave for each alphabet hence vary between 1 and 16. The code of the proposed algorithm can be written in any high level language. The number of occurrences of each alphabet is recorded and this is used to calculate the sine function's argument. The argument is such that the sine wave completes one cycle over the total occurrences of one alphabet. This is true for all 26 alphabets, irrespective of case. The watermark is embedded such that the colour of the text varies in the greyscale from 85 to 100 on a scale of 0-100. This ensures that the changes in the intensity of the black colour remain imperceptible. The text is read alphabet by alphabet and its colour property is changed according to the sine wave of the corresponding alphabet. In the case of special characters (such as exclamation marks, commas and full stops), the output colour is the same as the preceding alphabet. The document with the embedded watermark is converted to a pdf.

B. *Tamper Detection:*
At the receiver end, the raw text is extracted from the received pdf document. The embedding algorithm is again run on the raw text to generate a new watermarked document. This generated document is compared against the received document. Any mismatch reported will indicate that the received document has been tampered with. The two documents can be compared by converting them to images and subtracting them using some software like MATLAB. If the document has not been tampered with, a resultant black image will be obtained.

IV. PROPOSED ALGORITHM

EMBEDDING ALGORITHM

1. Read the raw text.
2. Run the SHA-1 over a pre-decided paragraph of the document to generate a hash.
3. 26 hexadecimal values out of the generated 40-hexadecimal hash are used as amplitudes for the sine waves. Let these values be $\{A_1, A^2.....A_{26}\}$.
4. The number of occurrences of each alphabet is recorded which then used for forming the argument for the sine function. Let these values be $\{N_1, N_2, ...N_{26}\}$.
5. Watermark is then embedded by changing the colour according to the sine wave of the corresponding alphabet is calculated in the following manner.
6. Let the color value of the m^{th} occurrence of the n^{th} alphabet be denoted by S_{nm} Then S_{nm} is given by the expression

$$\mathbf{S_{mn} = A_n * \sin(2\pi * m/N_n)}$$

Suppose, in a document, the letter "B" occurs a total of 5 times. Let the amplitude corresponding to it, as calculated by the hash function be 7. Then the greyscale value of the 3^{rd} (m=3) occurrence of "B" in the watermarked document will be calculated as:

$$S_{3,5} = 7 * \sin(2\pi * 3/5) = -4.1$$

Hence, the grey-scale value will be 93-4 =89. This indicates the integer colour value that "B" would attain on a grey scale of range 0-100.

DETECTION ALGORITHM

1. Extract the raw text from the received document.
2. Run the embedding algorithm again on it.
3. Convert the new watermarked pdf and the received pdf document to any image format.
4. Subtract the above two images (I_{m1} and I_{m2}) in order to compare them.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

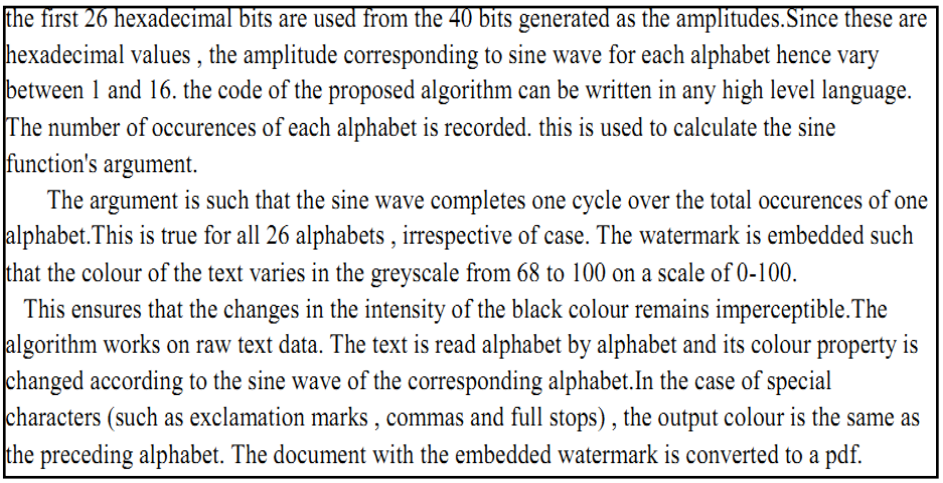
(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

5. The subtracted image is given by $\mathbf{I_{msub}= abs(I_{m1}-I_{m2})}$ and convert I_{msub} to greyscale format.
7. Plot the histogram.

V. EXPERIMENTAL RESULTS

The following figures demonstrate the experimental observations. **Figure 1** shows the raw text. The changes in the text document after embedding the algorithm are imperceptible, as seen in **Figure 2**. **Figure 3** shows the tampered document. The subsequent images show the detection process. **Figure 4** plots the histogram on comparing the un-tampered watermarked text obtained at the receiver's end against the text obtained by running the algorithm again on it. **Figure 5** illustrates the output on subtracting the images of the tampered watermarked text and the document obtained on running the algorithm again on the received tampered text. The histogram generated for **Figure 5** is shown in **Figure 6**. It can be seen from these results that if the document has been tampered with, then a resultant black image with grey patches is observed on subtraction. Otherwise, the resultant image will be completely black.


the first 26 hexadecimal bits are used from the 40 bits generated as the amplitudes.Since these are hexadecimal values , the amplitude corresponding to sine wave for each alphabet hence vary between 1 and 16. the code of the proposed algorithm can be written in any high level language. The number of occurences of each alphabet is recorded. this is used to calculate the sine function's argument.

The argument is such that the sine wave completes one cycle over the total occurences of one alphabet.This is true for all 26 alphabets , irrespective of case. The watermark is embedded such that the colour of the text varies in the greyscale from 68 to 100 on a scale of 0-100.

This ensures that the changes in the intensity of the black colour remains imperceptible.The algorithm works on raw text data. The text is read alphabet by alphabet and its colour property is changed according to the sine wave of the corresponding alphabet.In the case of special characters (such as exclamation marks , commas and full stops) , the output colour is the same as the preceding alphabet. The document with the embedded watermark is converted to a pdf.

Figure 1:Original Raw Text

The raw text shown in **Figure 1** when hashed with the SHA-1 algorithm (using a hash generator) generates the following 40-hexadecimal value:

"01c12fc24deaa65e4452335485854481603cd4bd"

The following table each alphabet along with its corresponding number of occurrences and the amplitudes calculated on the basis of the hash function. These values are then used as parameters {N_n and A_n} of the sine function to generate the instantaneous color value of the alphabets as explained in section IV. The total count and amplitudes for each alphabet are given in **Table 1**.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Alphabet	Count	Amplitude
A	81	1
B	18	2
C	49	13
D	30	2
E	133	3
F	18	16
G	15	13
H	59	3
I	49	5
J	0	14
K	5	15
L	37	11
M	24	11
N	45	7
O	56	6
P	25	15
Q	0	5
R	52	5
S	59	6
T	93	3
U	26	4
V	10	4
W	10	6
X	6	5
Y	7	9
Z	0	6

Table1: Alphabets with corresponding counts and calculated amplitudes

the first 26 hexadecimal bits are used from the 40 bits generated as the amplitudes.Since these are hexadecimal values , the amplitude corresponding to sine wave for each alphabet hence vary between 1 and 16. the code of the proposed algorithm can be written in any high level language. The number of occurences of each alphabet is recorded. this is used to calculate the sine function's argument. The argument is such that the sine wave completes one cycle over the total occurences of one alphabet.This is true for all 26 alphabets , irrespective of case. The watermark is embedded such that the colour of the text varies in the greyscale from 68 to 100 on a scale of 0-100. This ensures that the changes in the intensity of the black colour remains imperceptible.The algorithm works on raw text data. The text is read alphabet by alphabet and its colour property is changed according to the sine wave of the corresponding alphabet.In the case of special characters (such as exclamation marks , commas and full stops) , the output colour is the same as the preceding alphabet. The document with the embedded watermark is converted to a pdf.

Figure 2: Original Document after embedding Watermark

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Figure 3 shows the tampered document. The subsequent images show the detection process.

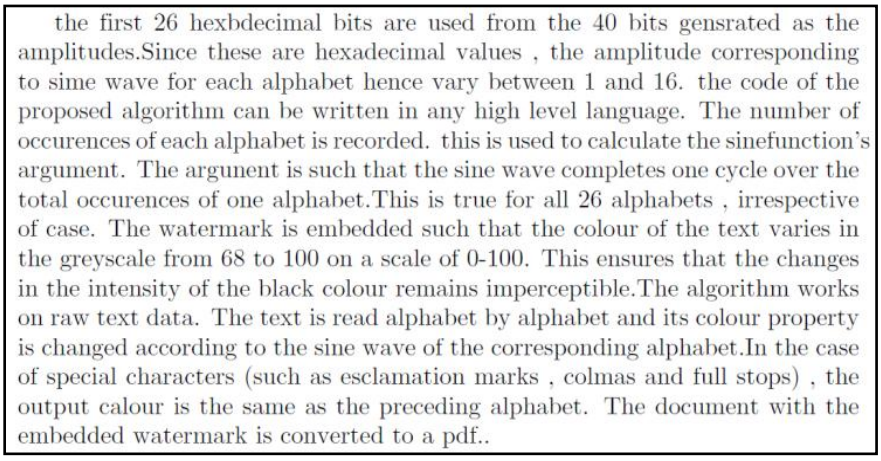

the first 26 hexbdecimal bits are used from the 40 bits gensrated as the amplitudes.Since these are hexadecimal values , the amplitude corresponding to sime wave for each alphabet hence vary between 1 and 16. the code of the proposed algorithm can be written in any high level language. The number of occurences of each alphabet is recorded. this is used to calculate the sinefunction's argument. The argunent is such that the sine wave completes one cycle over the total occurences of one alphabet.This is true for all 26 alphabets , irrespective of case. The watermark is embedded such that the colour of the text varies in the greyscale from 68 to 100 on a scale of 0-100. This ensures that the changes in the intensity of the black colour remains imperceptible.The algorithm works on raw text data. The text is read alphabet by alphabet and its colour property is changed according to the sine wave of the corresponding alphabet.In the case of special characters (such as esclamation marks , colmas and full stops) , the output calour is the same as the preceding alphabet. The document with the embedded watermark is converted to a pdf..

Figure 3: Tampered Document after running the Extraction Algorithm

When the document is untampered, the received pdf and the pdf generated after extraction are identical. Hence on subtraction, the resultant image is completely black. This causes the histogram to be concentrated around 0 as shown in **Figure 4**.

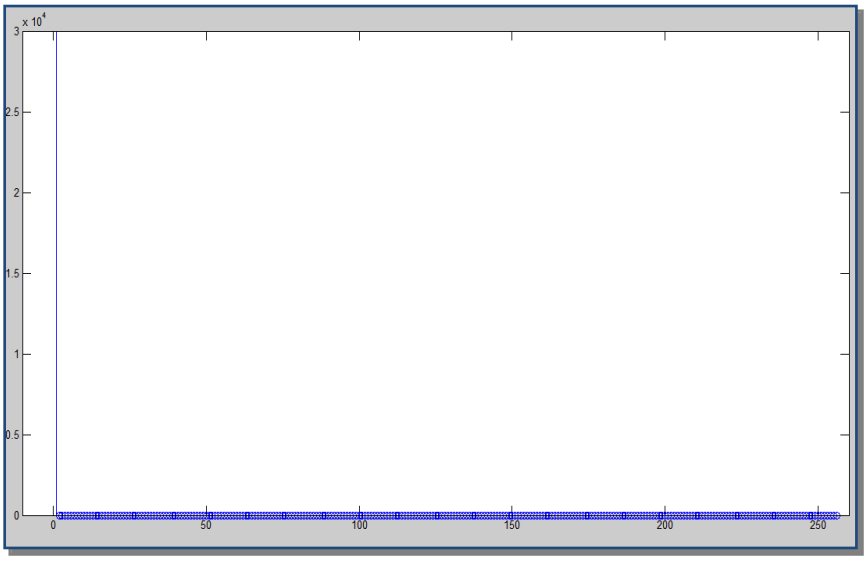


Figure 4: Histogram on comparing untampered watermarked document with the text obtained on running the algorithm again on it

Tampering of the document causes discrepancies in the parameters of the sine wave. Thus the resultant image has grey patches, shown in **Figure 5**.

.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

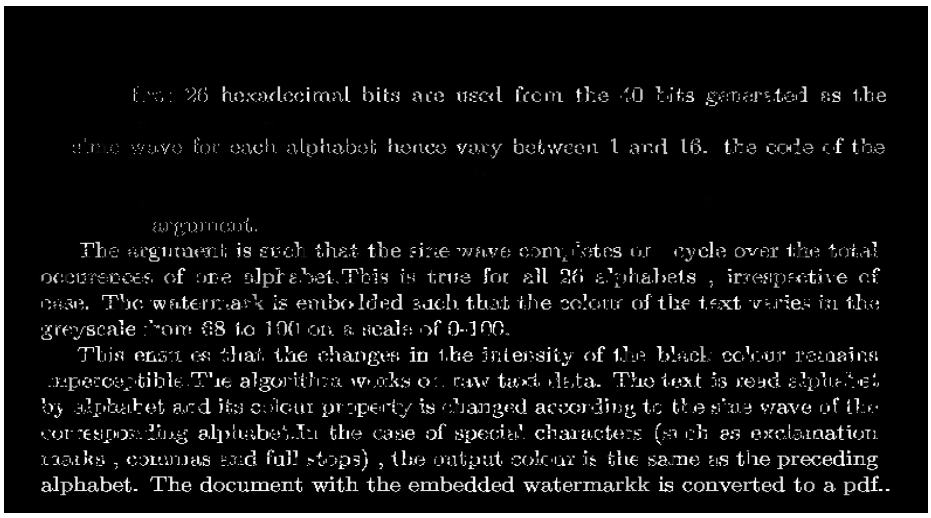


Figure 5: Subtraction of images of tampered text with that obtained in Figure 3

On plotting the histogram of this subtracted image, it is observed that it is no longer concentrated along zero but rather distributed along the whole greyscale as in **Figure 6**

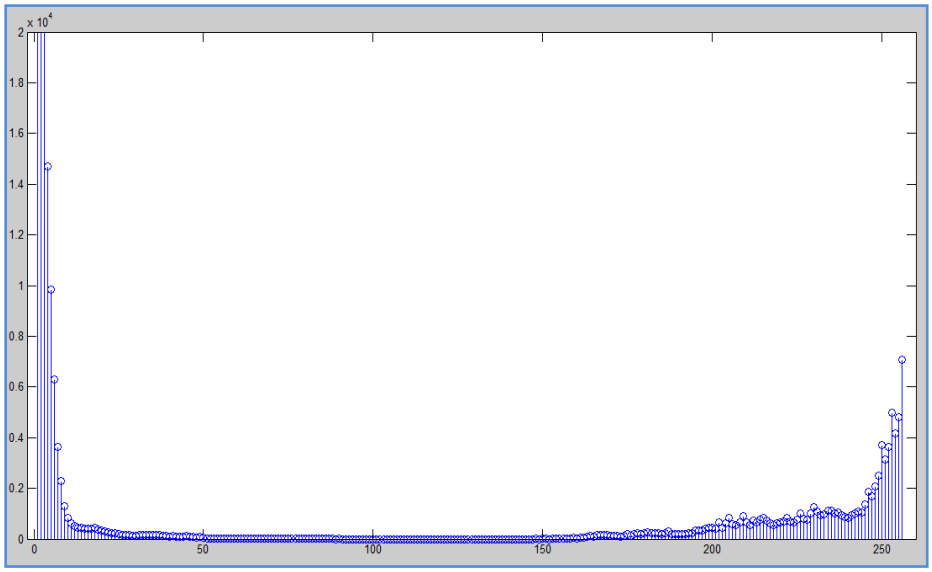


Figure 6: Histogram of Figure 5

VI. CONCLUSION AND FUTURE WORK

The algorithm implemented here is a novel watermarking scheme that is imperceptible and preserves the authenticity and integrity of the document. Using alphabet count as a variation parameter allows for an unprecedented amount of sensitivity of even small changes which is the strength of the proposed technique. The earlier works on watermarking for text authentication are not reliable in the case of random tampering attacks, especially when the amount of tampering is low. The algorithm hence is versatile, highly imperceptible and fragile. Future works may include extending the algorithm to watermark coloured documents and to detect changes in formatting, like indentation, of the document. With small changes, the algorithm may also be extended to other languages too.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

REFERENCES

[1] J. Brassil, S. Low, N.Maxemchuk, and L. O'Gorman, "Hiding Information in Document Images", Proceedings of the 29th Annual Conference on Information Sciences and Systems, Johns HopkinsUniversity, pp 482-489, March 1995.
[2] J. T. Brassil, S. Low, N. F. Maxemchuk, and L. O'Gorman, "Electronic Marking and Identification Techniques to Discourage Document Copying," IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp 1495-1504.
[3] N. F. Maxemchuk, S. H. Low, "Performance Comparison of Two Text Marking Methods," IEEE Journal of Selected Areas in Communications (JSAC) vol.16, no.4, pp 561-572, May 1998.
[4] N. F. Maxemchuk, "Electronic Document Distribution," AT&T Technical Journal, pp. 73-80. 6, September 1994.
[5] D. Huang and H. Yan, "Inter word distance changes represented by sine waves for watermarking text images," IEEE Trans. Circuits and Systems for Video Technology, Vol.11, No.12, pp.1237-1245, December, 2001.
[6] M. J. Atallah, C. McDonough, S. Nirenburg, and V. Raskin, "Natural Language Processing for Information Assurance and Security: An Overview and Implementations", Proceedings 9th ACM/SIG SAC New Security Paradigms Workshop, September, Cork, Ireland, pp. 51–65, 2000.
[7] M. J. Atallah, V. Raskin, M. C. Crogan, C. F. Hempelmann, F. Kerschbaum, D. Mohamed, and S.Naik, "Natural language watermarking: Design, analysis, and a proof-of-concept implementation", Proceedings of the Fourth Information
[8] PA. Hasan M. Meral, et al., "Syntactic tools for text watermarking", Hiding Workshop, vol. LNCS 2137, 9th SPIE Electronic Imaging Conf. 6505, Pittsburgh, 25-27th April 20011: Security, Steganography, and Watermarking of Multimedia Contents, San Jose , Jan. 2007.
[9] M. Atallah, C. McDonough, S. Nirenburg, and V. Raskin, "Natural Language Processing for Information Assurance and Security: An Overview and Implementations," Proceedings 9th ACM/SIGSAC New Security Paradigms Workshop, Cork, Ireland, pp. 51–65, September, 2000.
[10] U. Topkara, M. Topkara, M. J. Atallah, "The Hiding Virtues of Ambiguity: Quantifiably Resilient Watermarking of Natural Language Text through Synonym Substitutions".In Proceedings of ACM Multimedia and Security Conference, Geneva, 2006.
[11] Xingming Sun, Alex Jessey Asiimwe, "Noun-Verb Based Technique of Text Watermarking Using Recursive Decent Semantic Net Parsers", Lecture Notes in Computer Science (LNCS) 3612: 958-961, Springer Press, August 2005.
[12] Topkara, M., Topraka, U., Atallah, M.J., "Information hiding through errors: a confusing approach" In: Delp III, E.J., Wong, P.W. (Eds.), Security, Steganography, and watermarking of Multimedia Contents IX. Proceedings of SPIE-IS&T Electronic Imaging SPIE6505. pp. 65050V-1–65050V-12, 2007.
[13] B. Macq and O. Vybornova, "A method of text watermarking using presuppositions," in Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, January 2007.
[14] Peng Lu et al., "An optimized natural language watermarking algorithm based on TMR", on proceedings of 9th International Conference for Young Computer Scientists, 2009.
[15] Wiyada Yawai and Nualsawat Hiransakolwong, "Increase the Hiding-bit Capacity and Strength for Text Watermarking with the Line Intersection on Text Image" IEEE ICCM, vol. 1, pp. 427-433 , 2012.
[16] ZuneraJalil, Anwar M. Mirza, and Hajira Jabeen, "Word Length Based Zero-watermarking Algorithm for Tamper Detection in Text Documents," 2nd International Conference on Computer Engineering and Technology, vol. 6, pp. 378-382, 2010.
[17] Jalil Z., Farooq M., Zafar H., Sabir M., and Ashraf E. "Improved Zero Text Watermarking Algorithm against Meaning Preserving Attacks" World Academy of Science, Engineering and Technology, vol. 46, pp. 592-596 , 2010
[18] Xianmin Wei "Sine-wave-based Text Watermark for WORD Document", IEEE International Conference on Computer and Information Application, vol. 10, pp. 99-102,2010.
[19] Zunera Jalil, Anwar M. Mirza and Maria Sabir "Content based Zero-Watermarking Algorithm for Authentication of Text Documents", International Journal of Computer Science and Information Security, vol. 7, no. 2, pp. 212-217, February 2010.

BIOGRAPHY



Jeebananda Panda
Designation: Associate Professor
Qualification: ME
Specilization: Applied Electronics



Nishant Gupta
Designation:Student
Qualification: B.Tech
Department: Electronics and Communication Engineering



Parag Saxena
Designation:Student
Qualification: B.Tech
Department: Electronics and Communication Engineering

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015



Shubham Agrawal
Designation:Student
Qualification: B.Tech
Department: Electronics and Communication Engineering



Surabhi Jain
Designation:Student
Qualification: B.Tech
Department: Electronics and Communication Engineering



Asok Bhattacharyya
Designation:Professor
Qualification: PhD
Department: Electronics and Communication Engineering

IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE)
e-ISSN: 2278-1684,p-ISSN: 2320-334X, Volume 12, Issue 2 Ver. VI (Mar - Apr. 2015), PP 152-157
www.iosrjournals.org

Theoretical Investigation of Solar Energy Driven Combined Power and Refrigeration Cycle

Rahul Singh[1], Rajesh Kumar*[2] ,Supriya Vats[3], Pavan Kumar Yadav[4]
[1,2,3] *Department of Mechanical Engineering, Delhi Technological University (Government of NCT of Delhi), Bawana Road, Delhi-110042,* [4] *ISM Dhanbad Jharkand-826004, India*

Abstract; *This investigation is done for energy and exergy analysis of a combined power refrigeration cycle using oil as the heat transfer medium. This cycle is an integration of Rankine cycle for power production and ejector refrigeration cycle for cold production. The effects of parameters like; steam temperature, and the evaporator temperature of ejector have been observed on first and second law performance. The first law efficiency of solar driven combined cycle is found to be is 20% while second law efficiency is 11%.*
Keywords: *Solar energy, ejector, first law efficiency, second law efficiency*.

Nomenclature

A_h	aperture area of heliostat [m^{-2}]
q	solar radiation received per unit area [Wm^{-2}]
$\dot{E}$	exergy rate [kJ s^{-1}]
$\dot{Q}$	energy rate [kJ s^{-1}]
$\Delta\dot{E}$	exergy change [kJ s^{-1}]
RC	Rankine cycle
ERC	ejector refrigeration cycle
μ	entrainment ratio
η	efficiency[%]

Subscript

E	evaporator
C	condenser
CR	central receiver
D	destruction
HRVG	heat recovery vapor generator
EJE	ejector
d	diffuser
n	nozzle
m	mixing chamber
pf	primary flow
sf	secondary flow
n1	inlet of nozzle
n2	outlet of nozzle
1, 2, 3…….	state points in Fig.1

I. Introduction

Low temperature heat sources driven combined power and cooling thermodynamic cycle is proposed by Goswami (2000) has been under intense investigation. This cycle is the combination of Rankine power and vapor absorption refrigeration cycle operates on NH_3/H_2O mixture. On the other hand ejector refrigeration cycle has the advantages of simplicity in construction, low capital cost, high reliability, silent operation and very low maintenance cost. The refrigerants applied to the ejector refrigeration cycles, has almost zero ozone depletion potential and moderate global warming potential. Very few investigations are reported in the literature on thermodynamic evaluation of ejector refrigeration cycles. Sun and Eames (1996) described a simulation model for ejector refrigeration system working with R123 as an alternative for R11. Their results showed that R123 is a suitable replacement for R11 in space cooling applications.

Description of the proposed cycle: This cycle consists of six major components: heliostat field, central receiver, HRSG, turbine, evaporator, condenser, ejector, and pump and the schematic diagram is shown in Fig.1. The working fluid (4), which is water, is pumped to the HRSG (1) capable of producing the higher pressure of interest. Leaving the pump (1), the working fluid passes through a vertical stacked plate heat recovery steam generator (HRSG) where it generates steam from the heat source of 450^0C which oil that generate superheat steam at the temperature of 400^0C The superheated steam (2) passes through the turbine where electricity is produced with the help of generator. The bleed steam (5) from the turbine is mixed with the steam coming from the evaporator (7) and then goes to the top of the condenser (6) through a spray nozzle across the cooling element inside. The working fluid steam which is also working as refrigerant for cooling goes to the HRSG via pump (4). Water is filled in the evaporator (flash chamber) from where it is evaporated to the ejector (7) and cooling of the remaining water takes place in the flash chamber. The makeup water is entering at (8) as shown in figure 1

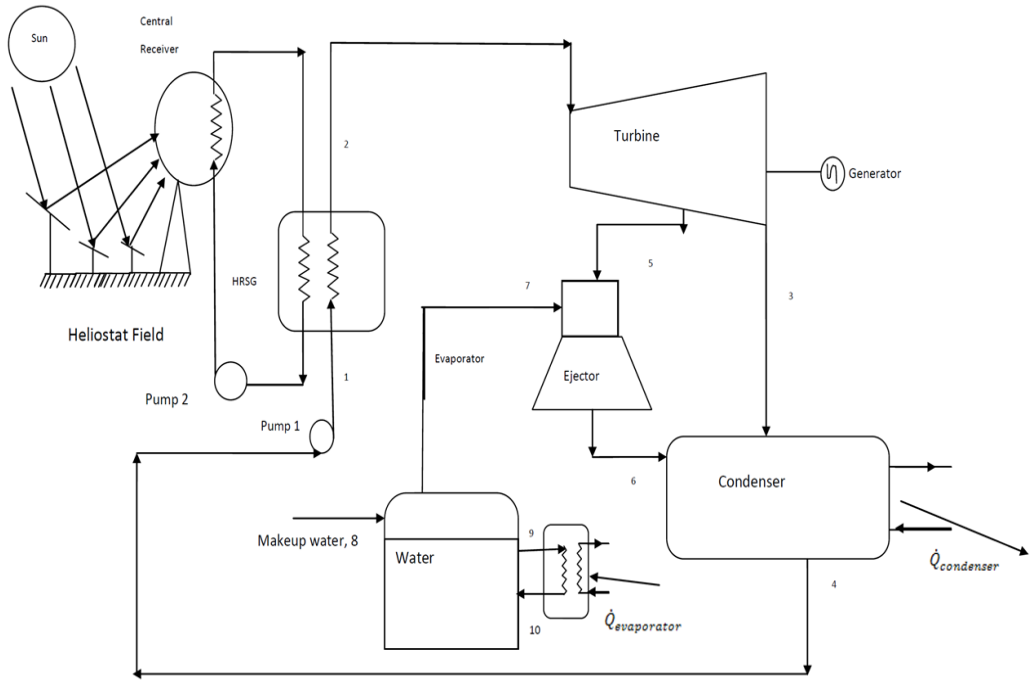


Figure 1. Schematic Diagram of the proposed solar thermal driven combined power and cooling system

The following assumptions have been made for the analysis of the proposed cycle;

(1) The components of the cycle are in steady state, and pressure drop in pipes and heat losses to the environment in the HRVG, turbine, condensers and evaporators are neglected.
(2) The flow through the throttle valve is isenthalpic.
(3) The condenser outlet state is saturated liquid.
(4) The evaporator outlet state is saturated vapor.
(5) Only physical exergies are used for the solar heat source and vapor flows.
(6) Kinetic, potential and chemical exergies of the substances are neglected.

For the analysis, the specifications of the combined RC, ERC,cycle are given in Table 1

Table 1: Main parameters considered for the analysis

Environment Temperature(^{0}C)	15
Turbine inlet pressure range (MPa)	4
Hot oil outlet temperature (^{0}C)	450(^{0}C)
Steam temperature(^{0}C)	400
Solar radiation received per unit area(Wm^{-2})	850
Apparent Sun temperature(K)	4500
Heliostat aperture area(m^2)	3000

Turbine back pressure range (MPa)	0.0074
Turbine isentropic efficiency (%)	85
Condenser temperature (^{0}C)	40
Hot oil mass flow rate (kg s^{-1})	8.0
Ejector evaporator temperature range (^{0}C)	4-10
Pump isentropic efficiency (%)	70
Nozzle efficiency (%)	90
Mixing chamber efficiency (%)	85
Diffuser efficiency (%)	85
First law efficiency of heliostat field (%)	75
First law efficiency of central receiver (%)	90
Second law efficiency of heliostat field (%)	75
Second law efficiency of central receiver (%)	30

Thermodynamic analysis: The basic principle of the model was introduced by Keenan et al. (1950) based on gas dynamics and formulated by Huang et al. (1999) and Ouzzane et al.(2003). The formulation and assumption of entrainment ratio is based on mass, momentum and energy equations which is recently developed by Dai et al.(2009) and may be reported as

$$\mu = \sqrt{\eta_n \eta_m \eta_d (h_{pf,n1} - h_{pf,n2,s'})/(h_{mf,d,s'} - h_{mf,m})} - 1 \quad (1)$$

The efficiencies of nozzle, mixing chamber, and diffuser are reported in Table 1 and the required enthalpy values at various state points of the ejector cycle for a given refrigerant are taken from REFPROP 6.01(1998). The ejector mainly consists of three section that is nozzle, mixing and diffuser section. In the nozzle section, the energy conservation equation for the adiabatic and steady primary flow is given as:

$$\dot{m}_{pf} h_{pf,n2} + \frac{\dot{m}_{pf} u_{pf,n2}^2}{2} = \dot{m}_{pf} h_{pf,n1} + \frac{\dot{m}_{pf} u_{pf,n1}^2}{2} \quad (2)$$

The nozzle efficiency may be defined as:

$$\eta_n = \frac{h_{pf,n1} - h_{pf,n2}}{h_{pf,n1} - h_{pf,n2,s'}} \quad (3)$$

In the mixing section, the momentum conservation equation is given as:

$$\dot{m}_{pf} u_{pf,n2} + \dot{m}_{sf} u_{sf,n2} = (\dot{m}_{pf} + \dot{m}_{sf}) u_{mf,m,s'} \quad (4)$$

In the diffuser section, the energy equation is given as:

$$\frac{1}{2}\left(u_{mf,m}^2 - u_{mf,d,s'}^2\right) = h_{mf,d,s'} - h_{mf,m} \quad (5)$$

The diffuser efficiency is given as:

$$\eta_d = \frac{h_{mf,d,s'} - h_{mf,m}}{h_{mf,d} - h_{mf,m}} \quad (6)$$

Exergy analysis determines the system performance based on exergy, which is defined as the maximum possible reversible work obtainable in bringing the state of the system to equilibrium with that of the environment. In the absence of magnetic, surface tension effect, nuclear, electrical, and considering the system is at rest relative to the environment, the total exergy associated with the work obtainable by bringing a stream of matter from its initial state to a state that is in thermal and mechanical equilibrium with the environment. Mathematically,

$$\dot{E} = \dot{m}[(h - h_0) - T_0(s - s_0)]$$

(7)

According to Gouy-Stodola theorem, the exergy destruction and entropy generation are related as

$$\dot{E}_D = T_0 \dot{S}_{gen} \quad (8)$$

First law Efficiency (η_I): It can be defined as the ratio of the desired effect to the thermal energy of solar input $\dot{Q}_{solar}$.

The first law efficiency of the triple effect cooling cycle is given by

$$\eta_I = \frac{\dot{Q}_{E1} + \dot{W}}{\dot{Q}_{Solar}} \quad (9)$$

The basic equation obtained from the law of conservation of energy in the components of RC, and ERC are written as follows:

For Heliostat: A part of thermal energy received by heliostat is delivered to the central receiver and rest is lost to the environment

$$\dot{Q}_{Solar} = A_h q$$

Where, A_h and q are the aperture area and solar radiation per unit area

$$\dot{Q}_{Solar} = \dot{Q}_{CR} + \dot{Q}_{lost,heliostat} \quad (11)$$

$$\eta_{I,heliostat} = \frac{\dot{Q}_{CR}}{\dot{Q}_{Solar}} \quad (12)$$

For Central Receiver (CR): A part of thermal energy received by central receiver is absorbed by oil and rest is lost to the environment

$$\dot{Q}_{CR} = \dot{Q}_{oil} + \dot{Q}_{lost,CR} = \dot{m}_{oil}(h_1 - h_{12}) + \dot{Q}_{lost,CR} \quad (13)$$

So, $\eta_{I,CR} = \frac{\dot{Q}_{oil}}{\dot{Q}_{CR}}$ (14)

Second law Efficiency (η_{II}) : The amount of exergy supplied in the product to the amount of exergy associated with the fuel is more accurate measure of the thermodynamic performance of the system which is defined as the ratio of exergy contained in the product to the exergy associated with the solar energy input and the second law efficiency of triple effect refrigeration cycle may be reported as

$$\eta_{II} = \frac{\Delta\dot{E}_{E1} + \dot{W}}{\dot{E}_{Solar}} \quad (15)$$

Where, $\dot{E}_{Solar}$ is incoming exergy associate with solar radiation falling on heliostat, $\Delta\dot{E}_{E1}$ is the change in exergy at ejector evaporator of ERC, $\dot{W}$ is the work output.

$$\dot{E}_{Solar} = \dot{Q}_{Solar}\left(1 - \frac{T_0}{T_S}\right) \quad (16)$$

Where, T_s the apparent sun temperature = 4500K
For Heliostat: A part of exergy received by heliostat is delivered to the central receiver and rest is lost to the environment (irreversibility),

$$\dot{E}_{Solar} = \dot{E}_{CR} + \dot{E}_{lost,heliostat} \quad (17)$$

$$\text{So, } \eta_{II,heliostat} = \frac{\dot{E}_{CR}}{\dot{E}_{Solar}} \quad (18)$$

For Central Receiver (CR): A part of exergy received by central receiver is absorbed by oil and rest is lost to the environment

$$\dot{E}_{CR} = \dot{E}_{oil} + \dot{E}_{lost,CR} \quad (19)$$

$$\text{So, } \eta_{II,CR} = \frac{\dot{E}_{oil}}{\dot{E}_{CR}} \quad (20)$$

II. Results and discussion

Energy and exergy analysis has been done to find out the effect of some influenced parameters on the performance of the solar driven combined power and refrigeration system. Following parameters have been chosen in the typical range of its operation; Motive steam pressure 4-7 MPa, evaporator temperature of ERC(4^0C -10^0C). The parameter under consideration is varied over a given typical range while values of other parameters are kept constant at the level of base case values.
Fig.2 shows the variation of the first and second law efficiency of combined cycle with the change in evaporator temperature.

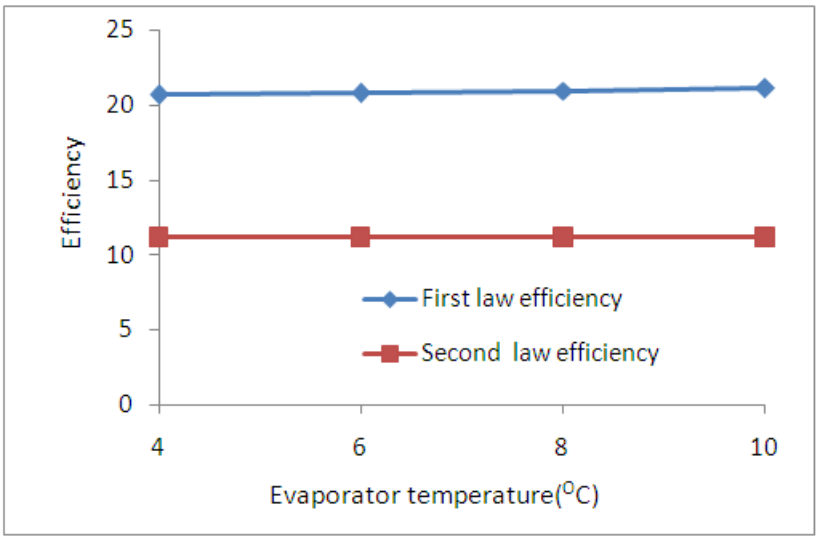


Figure 2: Variation of first and second law efficiency with the change in Evaporator temperature

And Fig.3 shows the variation of the first and second law efficiency of combined cycle with the change in motive steam pressure

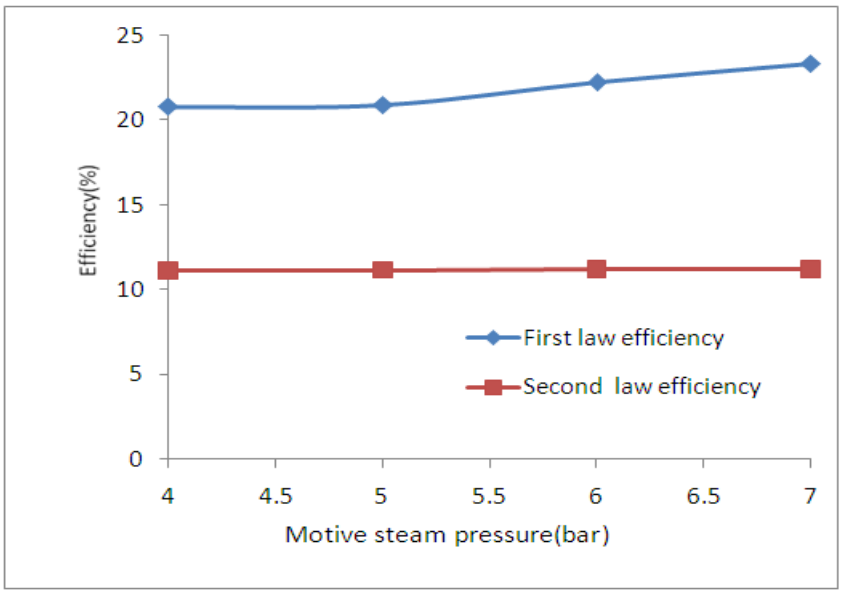


Figure 3: Variation of first and second law efficiency with the change in motive steam pressure

It is observed that first and second law efficiency increases with the increase in evaporator temperature and motive steam pressure. This is due to the fact that the refrigeration capacity of the cooling increases with the increase in evaporator temperature. The power produced by the turbine remains same so total effect of this is to increase the first and second law efficiency.

III. Conclusion

New solar driven combined power and cooling cycle is proposed for the production of cooling in the range of 4^0C to 10^0C. Energy and exergy methods are employed which enable us to develop a systematic approach that can be used to identify the sites of the real destructions/losses of valuable energy in thermal devices. The effect of design parameters were observed on energy and exergy performance of the proposed cycle. The conclusions of the present analysis can be summarized as follows:

- Energy efficiency of the cycle changes from about 20.5% to 21% with the change in evaporator temperature
- Exergy efficiency of the cycle changes from about 11.13% to 11.14% with the change in evaporator temperature
- Energy efficiency of the cycle changes from about 20.79% to 23.33% with the change in motive steam pressure
- Exergy efficiency of the cycle changes from about 11.14% to 11.22% with the change in motive steam pressure

References

[1]. F. Xu, D.Y. Goswami, S.S.Bhagat, 2000, A combined power/ cooling cycle, Energy, 25,233-246.
[2]. Da-Wen Sun, Ian W. Eames, 1996, Performance characteristics of HCFC-123 ejector refrigeration cycle, Int. J. of Energy Research, 20(10), 871-885.
[3]. Keenan H, Neumann EP, Lustwerk F. An investigation of ejector design by analysis and experiment. J Appl Mech, Transactions of the ASME 1950; **72**: 299-309.
[4]. Huang BJ, Chang JM, Wang CP, Petronko VA. A 1-D analysis of ejector performance. International Journal of Refrigeration 1999; **22**: 354-364.
[5]. Ouzzane M, Aidoun Z. Model development and numerical procedure for detailed ejector analysis and design. Applied Thermal Engineering 2003; **23**: 2337-2351.
[6]. Dai Y, Wang J, Gao L. Exergy analysis, parametric analysis and optimization for a novel combined power and ejector refrigeration cycle. Applied Thermal Engineering 2009; **29**: 1983-1990.
[7]. NIST Standard Reference Database 23, NIST Thermodynamic and Transport Properties of Refrigerants and Refrigerants Mixtures REFPROP, Version 6.01,1998.

Volume 3, Issue 1 (2015) 207-215 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Thermal Evaporation- Modeling and Microstructure Studies of Indium and Tin Deposition

Shailendra Kumar Gaur, R. S. Mishra
Department of Mechanical, Production & Industrial and Automobiles Engineering, Delhi Technological University, New Delhi, India

Article Info
Article history:
Received 20 February 2015
Received in revised form
25 February 2015
Accepted 28 February 2015
Available online 15 March 2015

Keywords
Physical Vapor Deposition,
Modeling,
Film Thickness Simulation,
Tin Evaporation,
Indium Evporation,
Thin Film Deposition

Abstract
The present work represents the modeling of nano scale tin & indium films by computing the film thickness, mass deposited on the substrate and mass transfer rate with time dependent model using BDF solver. Tin and indium is evaporated from a resistively heated evaporator source at a temperature of 1855 K and 1485 K respectively in a pressure (vacuum) of 100 Pa onto silicon surface held on a fixed surface. The film thickness varies between 144 nm to 164 nm for Tin and 164 nm to 183 nm for Indium across the sample after 60 sec of deposition, with radial symmetry about the midpoint of the source.
The film thickness as well as mass deposited at a point increases linearly with time. Since the angular distribution is of particular interest in this model, by increasing the integration resolution to a maximum value for ensuring the most accurate angular resolution when computing the flux. The surface temperature is required to specifying the temperature of the evaporating tin and indium source using constant elements for turn off the refinement in the post-processing settings. The SEM micrographs of tin and indium at different magnifications shows the 100nm to 1microns grain size along the grain boundaries. Similarly, XRD analysis with Kα (wavelength 1.541874) shows the peaks of intensity at different 2θ angles for different orientations of planes with polycrystalline structure.

1. Introduction

Physical methods produce the atoms that deposit on the substrate by evaporation and sputtering. Sometimes called vacuum deposition because the process is usually done in an evacuated chamber. PVD is used for metals while dielectrics can be deposited using specialized equipments. It rely on thermal energy supplied to the crucible or boat to evaporate atoms. Evaporated atoms travel through the evacuated space between the source and the sample and stick to the sample. Few, if any, chemical reactions occur due to low pressure can force a reaction by flowing a gas near the crucible. Surface reactions usually occur very rapidly and there is very little rearrangement of the surface atoms after sticking. Thickness uniformity and shadowing by surface topography, and step coverage are issues.Ali Moarrefzadeh et al. [1] discussed that Physical vapor deposition (PVD) includes a wide range of vacuum coating processes in which material is physically taken out from a source by evaporation or sputtering, transported through a vacuum or partial vacuum by the energy of the vapor particles, and condensed as a film on the surfaces of appropriately placed parts or substrates. A group of very versatile coating processes in which a material is converted to its vapor phases in a vacuum chamber and condensed onto a substrate surface as a very thin film (upto 1micron thickness). The deposition of thin film layers from the vapor phase is done through several methods. We study and analyze the physical vapor deposition (PVD) techniques and equipments that are in common use in the large scale production of coatings that find uses in the optical, display, decorative, tribological, and energy-generating/ saving industries. Evaporating materials are classified as dielectric compounds, metals, alloys, or mixtures. The same evaporant material can exhibit different optical, electrical, and mechanical properties depending on the deposition process. Titanium oxide is a unique example of a metal oxide compound that, depending on deposition process parameters, can be made into film layers that are: transparent, electrically conductive, chemically reactive to light and bio- agents, chemically inert, or exhibit spectrally selectively absorption. The dependent parameters are starting composition, oxidation state, and crystalline structure and packing density. PVD techniques used generally are basically two in nature: thermal evaporation by resistively heating or by using an electron- beam heating, and sputtering, a no thermal process. Alterations and accompaniments are made to the basic PVD techniques to permit different coating materials and substrate types to be included. Process additions designed to alter the growth nano-structure or composition of the film through control of the dependent variables listed above include bombardment of the growing film by high energy inert- or and reactive ions, substrate heating, atmosphere composition and partial pressure, rate, and vapor incidence angle. A further important variable contribution to the nucleation and self- assembling growth structure of the condensing atoms, that we have discussed frequently, is the condition both chemical and physical of the substrate surface (PVD process Deposited films are typically evaluated for visual defects, thickness, and adhesion. Visual defects such as bare spots, small voids, incorporated flakes, or debris can be observed with a stereo microscope having a magnification of 10 to 100 times. Film thickness

Corresponding Author,
E-mail address: professor_rsmishra@yahoo.co.in
All rights reserved: http://www.ijari.org

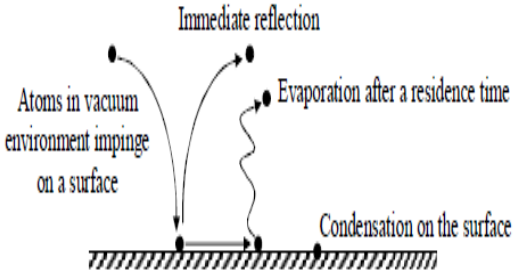


Fig: 1. Mechanism of Thermal evaporation process (ref. [1])

is generally measured by one of the following methods: Thickness of film as modeled can be simulated experimently with dektek surface profilometer with diamond size of 12.5µm radius stylus. Polished metallurgical micro sections are used to microscopically observe the coating thickness on various part surfaces. This method is the most direct way to determine thickness uniformity. Beta (high-energy electron) backscatter instruments are used to measure the film thickness nondestructively. This is an indirect method that requires calibration with a known standard; substantial errors can be made in measuring the film thickness on curved surfaces if care is not exercised. In thermal evaporation process, load the source material-to-be-deposited (evaporant) into the container (crucible). Then resistively heat the container (tungsten or molybdenum boat) containing source to high temperature i.e. evaporation temperature depending upon the material to be deposited. The source material evaporates and evaporant vapor transports to and impinges on the surface of the substrate. Evaporant condenses on and is adsorbed by the surface.

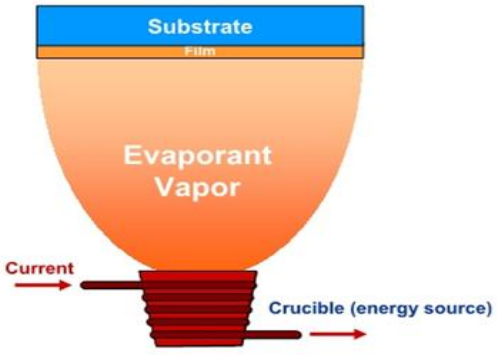


Fig: 2. Thermal evaporation process (ref. CIMS applied physics)

Further study shows that evaporation used in thin film deposition can be done by:

a) Thermal evaporation
b) E-beam evaporation

Thermal evaporation is very old, easy, reliable and good method of deposition while E-beam evaporation requires elaborate and complex system. Moreover high velocity electron creates surface defects and stress on the deposited surface**.** The present work has been performed to achieve the following objectives:

1. To find out the deposited film thickness on the silicon substrate and variation of the thickness across the surface of sample.
2. To find out the mass deposited on the silicon substrate and its variation with time
3. To find out the mass deposition rate on the silicon substrate with time
4. Micrograph and microsture analysis of the deposited film SEM and XRD for grain size along grain boundries and orientation of crystal planes for X-ray.

2. Literature Review

Thin tin oxide deposited surface is very useful in industries due to their properties such as n-type Semi-conductor character, high optical transmission in visible range , high reflectivity; in the infra-red as well as good chemical resistance. Thus they can be used to form transparent and chemically stable thermal barriers. C. R. Zamarreñoa et al. [3] discussed the fabrication of optical fiber refractometers based on indium tin oxide (ITO) coatings with good sensitivity in the visible spectral region. ITO thin-films have been deposited by using sputtering mechanism and employing a rotating mechanism that makes possible the fabrication of smooth homogeneous coatings onto the optical fiber core. Souad Laghrib et al. [4] discussed the vacuum evaporation of Tin followed by annealing with oxygen flow to obtain Tin oxide thin layers. He did a lot of experiments by varying the deposited Tin thickness as well as annealing conditions by varying annealing temperature with different duration and found that at high temperature annealing forms nanocrystalline tin oxide. Sea-Way Jan et al.[5] discussed the formation of indium tin oxide by first metal tin and indium deposition in oxygen in vacuum coating system maintain vacuum of 1x 10 -6torr by combination of rotary and diffusion pump. Furthur studied the deposited film through XRD, SEM, TEM, Hall measurement and found polycrystalline film with bcc structure. This film is used in front panels in liquid crystal and electroluminescence displays, transparent heating elements, high-efficiency ITO/GaAs and ITO/InP solar cells, blue light electroluminescence etc. D. Bruce Buchholz et al. [6] discussed the few nm thin layer of indium oxide by pulse layer deposition method and studied formation of crystalline and amorphous with variation of temperature and density variation. Alexandu C. Fechete et al. [7] discussed the indium oxide nanostructures in the form of nanobelts and nanorods with few microns length and 200nm width developed by thermal evaporation on silicon substrates with and without presence of gold catalysts in temperature range of 600- 900˚ C in which growth process involved are vapor-solid and vapor-liquid-solid and the nanobelts are appx.10-200 nm wide and few micrometers long at 750˚C when no catalyst while when Au catalyst is used, homogenous scattered nanorods obtained at 900˚C. C. A. Pan et al. [8] discussed that high quality and high conductive indium oxide films prepared by thermal evaporation from In and In2O3 source in vacuum chamber with low pressure of O2 of which properties comparable to tin doped indium oxide films. Its low resistivity is as a result of excellent electron mobility (~70cm2/V sec) though the

electron concentration is rather high which makes it a perfect for fabrication of semi-conductor device. Burstein shifts due to high electron density in conduction band also seen. B. Maniscalco et al. [9] discussed the thin film thickness measurement of Indium Tin Oxide made using coherence correlation interferometry. Optical transmission over the whole visible spectra with a steep decrease 380 nm. Indium Tin Oxide(ITO) is widely used in applications such as touch screens, LCD and LED displays, thin film photovoltaics. M.Trzcinski et al [10] discussed the alloy formation of ultrathin Ag and In layers deposited on W(100) starting at room temperature and finished after a small annealing at 600 K. The alloying process is originally geometrical and no major difference between 2D layered alloys and 3D bulk alloys can be found. Varieties of coatings can be deposited such as metals, alloys, ceramics and other inorganic compounds, and even certain polymers. Deposition can be done onto the varieties of substrates such as metals, glass, and plastics. The properties of atavistically deposited films depend strongly on:

- The material being deposited
- Substrate surface chemistry and morphology
- The surface preparation process
- The details of deposition process and the deposition Parameters :
- mean free path of evaporant
- source to substrate distance
- vapour pressure of evaporatnt
- sticking coefficient

Condensation and nucleation- Atoms that impinge on a surface in a vacuum environment may be reflected immediately, re-evaporate after a residence time, or condense on the surface shown in figure

1. Mean free path is the minimum distance between two successive collisions and is given by formula depending upon the vacuum in the chamber:

$$\lambda = \frac{k_B T_c}{\sqrt{2}\pi \cdot P \cdot d^2}$$

Where T– usually room temp of chamber P- vacuum in chamber based on this mean free path source to substrate distance is optimized for constant flux and better deposited film quality. Vapor pressure varies depending upon the source temperature and is different for different materials, for suitability standard plot [11] is provided from which depending upon source material vapor pressure is selected. Sticking coefficient is defined as the ratio of the condensing atoms to impinging atoms. If the atoms do not immediately react with the surface, they will have some degree of surface mobility over the surface before they condense. Re-evaporation is a function of bonding energy between the adatom and the surface, the surface temperature, and the flux of mobile adatoms e.g. the deposition of cadmium on a steel surface having a temperature greater than about 200 ˚C will result in total re-evaporation of the cadmium. The energy of the atom, atom-surface interaction (chemical bonding), and the temperature of the surface influence the capability of an atom on a surface. The mobility on a surface can change because of variations in chemistry or crystallography. The various crystallographic plates of a surface have different surface free energies that influence the surface diffusion. Atoms condense on a surface by losing energy. They lose energy by:

- Forming and breaking chemical bonds with the substrate atoms.
- Finding preferential nucleation sites (lattice defects, atoms steps, and impurities)
- Interacting with another diffusing surface atoms (same species)
- Colliding or reacting with adsorbed surface species.

Atoms form nuclei after condensation. Homogenous nucleation is said to occur if the surface is of same material as the deposition atoms and if they are different materials, the process is called heterogeneous. In semiconductor field, heterogeneous nucleation forms hetero- junctions. G.H. Gilmer [2] et al discussed that **t**hree types of nucleation mechanisms have been identified; they differ according to nature of interaction between the posited atoms and the substrate material. Frank-Van der Merwe mechanism leading to a monolayer-by-monolayer growth (layer growth; ideal epitaxy). Volmer-Weber (V-W) mechanism, characterized by a three-dimensional nucleation and growth (island growth). Stranski-Krastanov (S-K) mechanism, where an altered surface layer is formed by reaction with the deposited material to generate a strained or pseuomorphic structure, followed by nucleation on this layer (Layer + island growth) nuclei coalescence and agglomeration The nuclei develop by collecting atoms that diffuse over the surface. Isolated nuclei grow laterally and vertically on the surface to form a continuous film and to form a continuous film less amount of material is needed if nucleation density is high. The principal growth mode of nuclei may be: laterally over the substrate surface (wetting growth) such as gold on copper and chromium, iron on W-O surfaces, and titanium on SiO2 the nuclei may prefer to grow in a vertical mode (dewetting growth) such as nickel and copper on W-O surfaces, and gold on carbon, Al2O3, and SiO2. Growth and coalescence of the nuclei can leave interfacial voids or structural discontinuities at the interface, particularly if there is no chemical interaction between the nuclei and substrate material and dewetting growth occurs. Though a lot of research work has been done on the evolution of tin and indium film deposited on nanoscale or higher thickness by different deposition process through different heating sources (resistively, sputtering, E-beam, magnetron, radio frequency) in different baskets/crucibles at different temperatures and modeled mostly with direct simulation monte carlo. Also these methods solve in the volumes of the modeled and geometries. Number densities found to be not accurate and precise. Ineligible simulation for the accurate modeling of low pressure, low velocity gas flows in complex geometries. These studies subjected to statistical scatter. Moreover, these could not completely explained free molecular flow interface and applied dsmc computes the trajectories of large numbers of randomized particles through the system, but introduces statistical noise to the modeling process and also the method is slower.

3. Experimental Set-up

The island films were prepared using a laboratory thermal evaporation setup working at residual vacuum of (2.5 - 4) × 10¯6mbar. The deposition setup was equipped with the two-stage vacuum system based on the diffusion

pump (oil based), turbo molecular or cry pump along with rotary vane pump or rotary screw pump (oil free). In case of oil based pumps the ultimate vacuum achieved in the system depends upon the vapour pressure of the oil used in pumps, generally Silicone based oils are used to achieve vacuum of1.0 x E-6 order in the chamber. Sometimes high discharge Ar gas plasma substrate cleaning can be done at vacuum of 5 x E-3 mbar, if needed. The films were deposited on the substrates cleaned in the ultrasound bath in isopropyl alcohol and drained by a compressed air flux. During the deposition, all substrates were kept at room temperature i.e

Table: 1. Typical Boat /Crucible Material

Refractory Metals		
Material	**Melting Point (°C)**	**Temperature for 10-mtorr Vapor Pressure (P_e) (°C)**
Tungsten (W)	3380	3230
Tantalum (Ta)	3000	3060
Molybdenum (Mo)	2620	2530
Refractory Ceramics		
Graphitic Carbon (C)	3799	2600
Alumina (Al_2O_3)	2030	1900
Boron Nitride (BN)	2500	1600

4. Result Analysis

General description of the model is presented in fig. 1 where low pressure gas flow in vacuum system i.e molecular flow and it can be seen that the model inputs are ambient temperature, evaporant temperature, vapour pressure, molecular weight, density of gold, neglecting the substrate surface temperature. The Free Molecular Flow interface uses the angular coefficient method to model flows with Knudsen numbers Kn >10. This physics interface avoids solving the physics in the volumes of the modeled geometries, and requires meshing only of the surfaces. Completely diffuse scattering (total accommodation) and emission are assumed at all surfaces in the geometry, and flow is computed by integrating the flux arriving at a surface from all other surfaces in its line-of-sight. This means that the dependent variables exist only on the surfaces of the geometry, and the solution process is much faster than the DSMC method. Furthermore, it is not subject to statistical scatter. Number densities are reconstructed using a method included in the Free Molecular Flow interface. The Molecular Flow Module is designed to offer previously unavailable simulation capabilities for the accurate modeling of low pressure, low velocity gas flows in complex geometries. It is ideal for the simulation of vacuum systems including those used in semiconductor processing, particle accelerators and mass spectrometers. Small channel applications (e.g. shale gas exploration and flow in nanoporous materials) can also be addressed. The Molecular Flow Module uses a fast angular coefficient method to simulate steady-state free molecular flows. We can model isothermal and nonisothermal molecular flows, and automatically calculate the heat flux contribution from the gas molecules. The discrete velocity method is also included in the module for the simulation of transitional flows. Historically, flows in this regime have been modeled by the direct simulation Monte Carlo (DSMC) method. This computes the trajectories of large numbers of randomized particles through the system, but introduces statistical noise to the modeling process. For low velocity flows, such as those encountered in vacuum systems, the noise introduced by DSMC renders the simulations unfeasible. COMSOL uses alternative approaches: employing a discrete velocity method for transitional flows (using a Lattice Boltzmann velocity quadrature) and the angular coefficient method for molecular flows. Using the input parameters, the model computes the thickness of gold deposited, mass deposited on the surface of substrate, mass transfer rate and mass loss by using COMSOL software. The following wall conditions are inbuilt in COMSOL software:

wall, out gassing wall, adsorption/desorption, deposition. In adsorption/desorption boundry condition, sticking coefficient can be defined along with other condition. Tin and Indium is placed in resistively heated Tungstun boat which is having very high melting point of 3420^0C. Substrate is one quarter of a 4″ wafer mounted on stationary support on top of tungsten boat depending upon the mean free path and Langumuire-Kundsen relation. A screen is placed to cover the substrate, if more than one source is used.

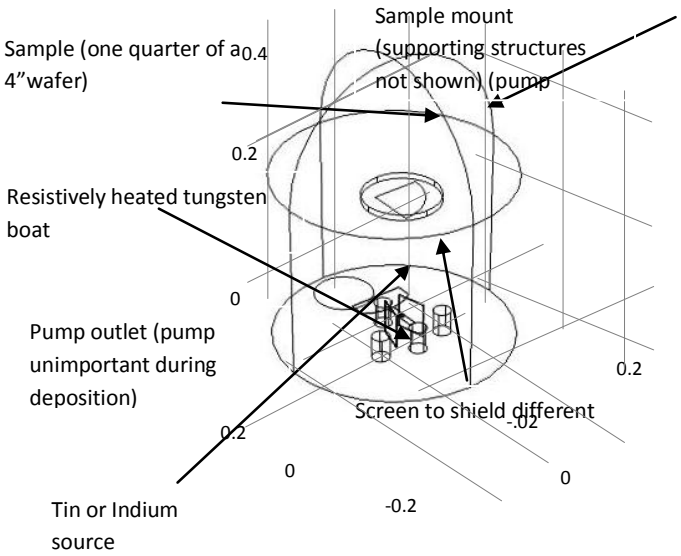


Fig: 3. Model Geometry Various Components of evaporator

The basic governing equation for depositing molecular flux on the surface is Langmuire-Knudsen Relation which shows the mass deposition rate per unit area of source surface:

$$R_m = C_m\left(\frac{M}{T}\right)^{\frac{1}{2}}\cos\theta\cos\varphi\frac{1}{r^2}\left(P_e(T)-P\right)$$

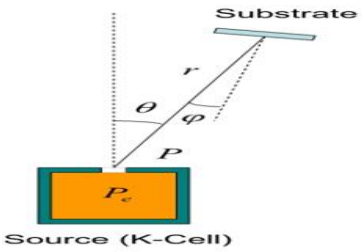


Fig: 4. Source to Substrate Geometry

$C_m = 1.85x10^{-2}$
r: source-substrate distance (cm)
T: source temperature (K)
P_e: evaporant vapor pressure (torr), function of T
P: chamber pressure (torr)
M: evaporant gram-molecular mass (g)

Uniformity on a Flat Surface

Consider the deposition rate difference between wafer center and edge:

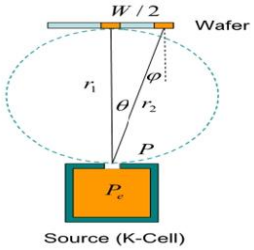


Fig: 5. Source to Flat Surface Geometry

$$R_1 \propto \frac{1}{r_1^2}$$

$$R_2 \propto \frac{1}{r_2^2}\cos^2\theta = \frac{r_1^2}{r_2^4}$$

Define Uniformity:

$$\sigma(\%) = \frac{R_1 - R_2}{R_1}(\%)$$

$$\sigma = 1 - \left(1 + \left(\frac{W}{2r_1}\right)^2\right)^{-2} \approx \frac{W^2}{2r_1^2} \quad \text{or} \quad \boxed{\frac{W}{r_1} = \sqrt{2\sigma}}$$

5. Results & Discussion

In table 2 & 3 various input parameters required for the tin and indium models are given to compute the molecular flow in vacuum system. Generally, the ambient temperature is fixed based on the experimental setup or experimental conditions but the evaporation temperature can be varied as per the thickness or mass transfer to be deposited on the substrate. This evaporation temperature can be increased or decreased depending considering the vapour pressure, to make fast the mass transfer keeping the good adhesion with uniformity on the substrate surface.

In table 2 &3 various input parameters required for the tin and indium models are given to compute the molecular flow in vacuum system. Generally, the ambient temperature is fixed based on the experimental setup or experimental conditions but the evaporation temperature can be varied as per the thickness or mass transfer to be deposited on the substrate. This evaporation temperature can be increased or decreased depending considering the vapour pressure, to make fast the mass transfer keeping the good adhesion with uniformity on the substrate surface.

Table: 2. Input Parameter for the tin Model

Name	Expression	Value	Description
Tamb	293.15[K]	293.2 K	Ambient temperature
Tevap	1855[K]	1855 K	Evaporation temperature
pvap	100[Pa]	100.00 Pa	Vapor pressure of Tin
Mn0	118.71[g/mol]	0.11871kg/mol	Molecular weight of Tin
Rho0	7.36 g/cm^3	7.36E3 kg/m³	Density of Tin

Table: 3. Input Parameters for the Indium Model

Name	Expression	Value	Description
Tamb	293.15[K]	293.2 K	Ambient temperature
Tevap	1485 [K]	1485 K	Evaporation temperature
pvap	100[Pa]	100.00 Pa	Vapor pressure of indium
Mn0	114.818[g/mol]	0.1148 kg/mol	Molecular weight of indium
Rho0	7.31 g/cm^3	7.31E3 kg/m³	Density of indium

Figure 6 shows the flux of tin molecules on the surfaces of the model. This constant flux determines the thickness of Tin deposited shown in figure. In figure 7, it is obvious the tin film thickness is maximum in the centre of substrate positioned directly on vertically centre of tungsten boat/source, greater than 164 nm after 60 seconds of thermal evaporation process. This thickness decreases radialy outwards surface of the substrate to 144 nm. In Figure 8 based the molecular flux deposited on substrate, graph shows the tin deposited thickness after 60 seconds from initial condition of zero second. It is clear from graph that the film thickness varies linearly with time. In figure 9 graph is plotted with time for mass (kg x E-6) deposited on the surface of substrate and shows that mass deposited on substrate increases with time as in the case of figure 8 since the depositing thickness with time adds mass also on the substrate. In figure 10 mass transfer rate (kg x E-8/sec) is plotted for 60 seconds and is a straight constant abscissa parallel line.

Time = 60 Surface Film thickness (nm)

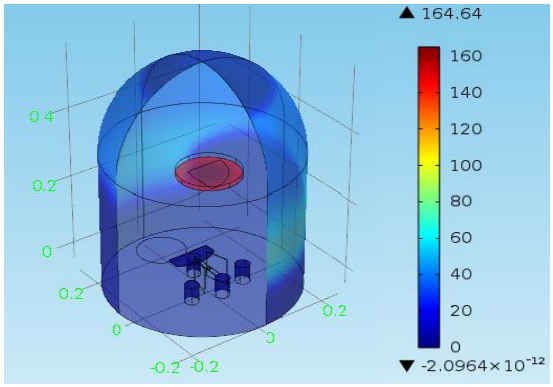


Fig: 6. Tin film Deposited Thickness

Time = 60 sec Surface: film thickness (nm)

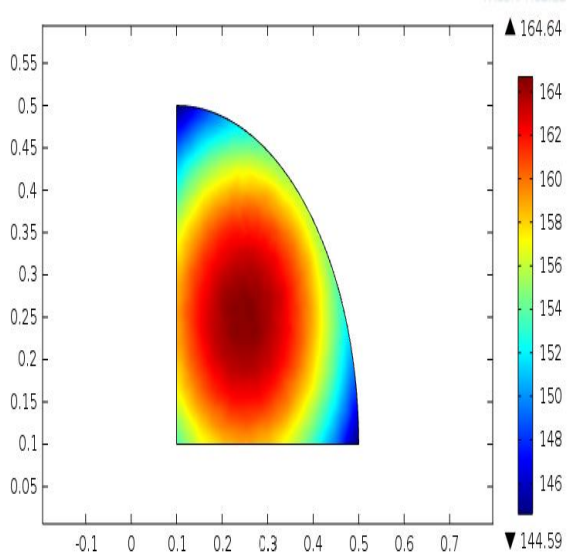


Fig: 7. Variation of Deposited tin Film Thickness Across the Surface

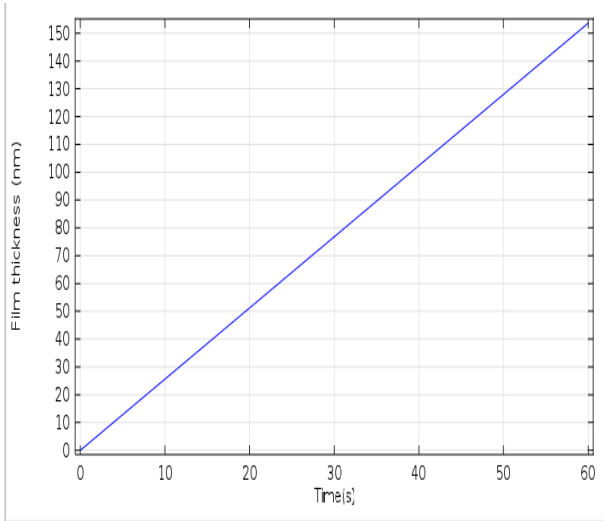


Fig: 8. Film Thickness vs Time at a Point on the Corner of the Sample

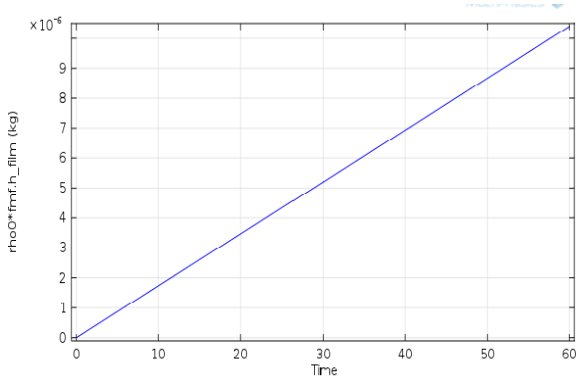


Fig: 9. Mass of tin deposited on the surface

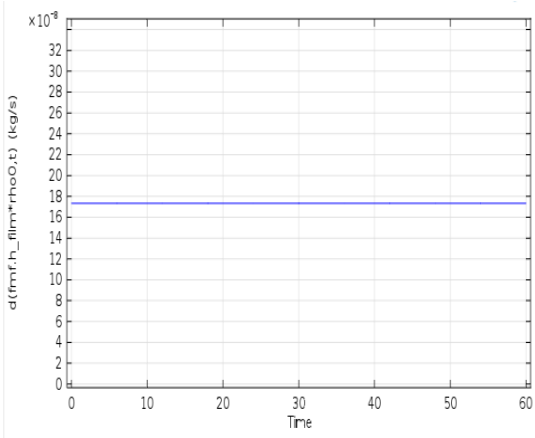


Fig: 10. Variation of Mass Transfer Rate of tin with Time

Figure 11 shows the deposition of indium at 100 Pa with evaporation temperature of 1485 K. The maximum film thickness achieved is 183 nm, this film thickness may vary by varying the evaporation temperature with time and vapor pressure. Similarly, the indium thin film thickness varies from 160 to

183 nm from the centre of substrate to outer surface and it is maximum at the centre as shown in figure 12.

Time = 60 sec Surface film thickness (nm)

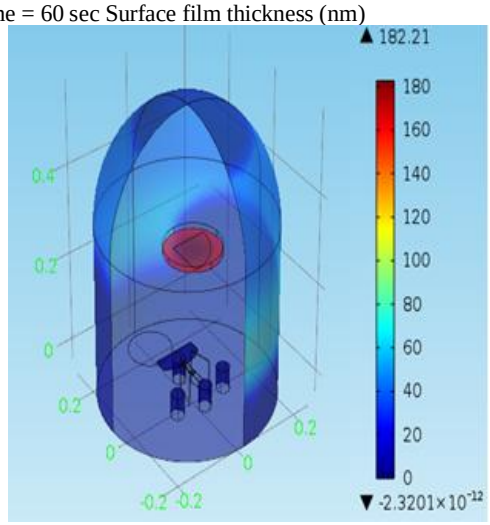


Fig: 11. Indium Film Deposited Thickness

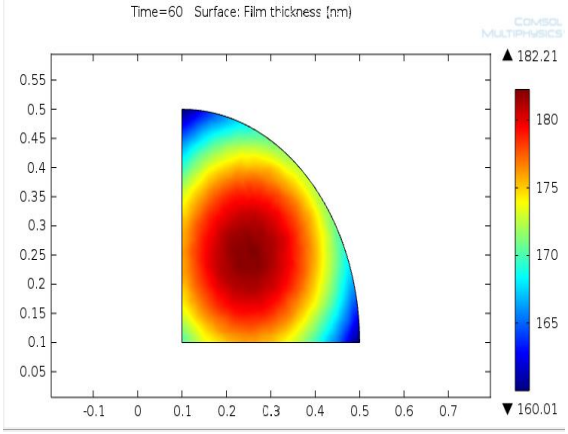


Fig: 12. Variation of Deposited Indium Film Thickness Across the Surface

Figure 13 shows the variation of thickness at a point on the surface and its variation with time i.e after 60 sec from zero to 183nm. This plot is taken at present at the edge of substrate and similarly by selecting the another point on the surface, we can find similar plot.

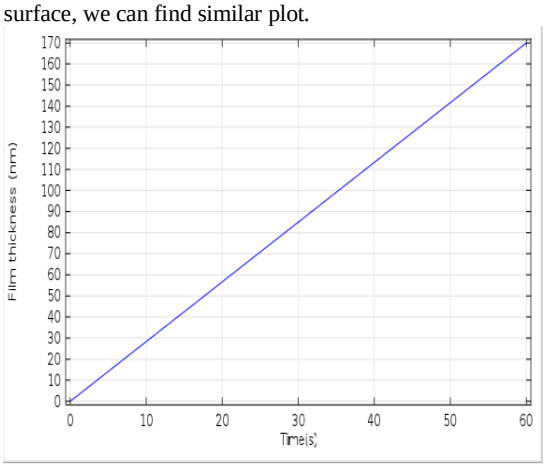


Fig: 13. Film Thickness vs Time at a Point on the Corner of the Sample

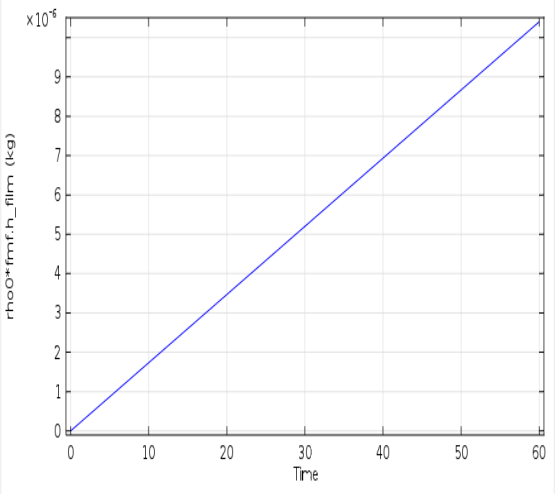


Fig: 14. Mass of Indium Deposited on the Surface

Figure 14 shows the mass (kg x E- 6) of indium deposited on the surface of substrate and shows that as mass deposited on substrate increases with time. Now as in the case of figure 13 where the depositing thickness with time adds, mass also on the substrate deposited increases.

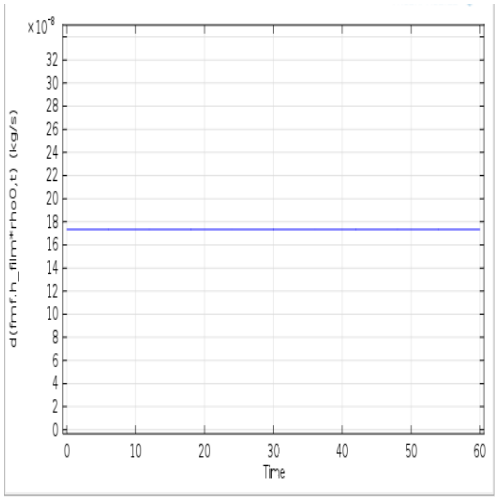


Fig: 15. Variation of Mass Transfer Rate of Indium with Time

Figure 15 shows mass transfer rate (kg x E-8/sec) is plotted for 60 seconds and is a straight constant abscissa parallel line. Only the flux is required to compute the deposition rate, but in this instance, since most of the computational time is used to compute the view factors, solving the time dependent problem adds little additional time to the solution process. Using a time dependent model also allows for more advanced extensions of the model, for example, re-evaporation of gold from hot surface close to the evaporative source could be included. Definitely there is loss of source material during thermal evaporation since only a fraction of the material is deposited on the surface while we are evaporating a large quantity of it. The SEM micrographs shown in figure 16, 17 are at 5KX, 10KX magnifications while fig. 18, 19 are at 25KX magnification, 20KV Extra High Tension (EHT), working distance 10.6mm to go through better study and analysis of grain size and structure. The SEM revealed the growth of 200nm-1µm grains size.

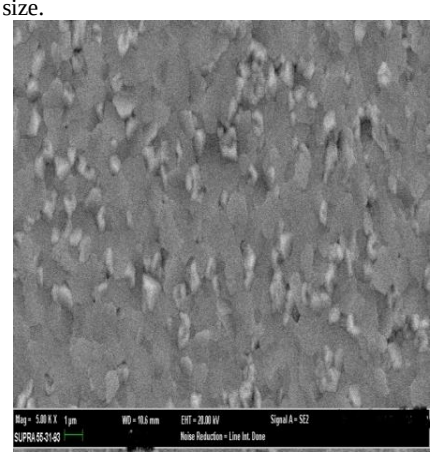


Fig: 16. SEM Micrographs of tin at 5X Magnification

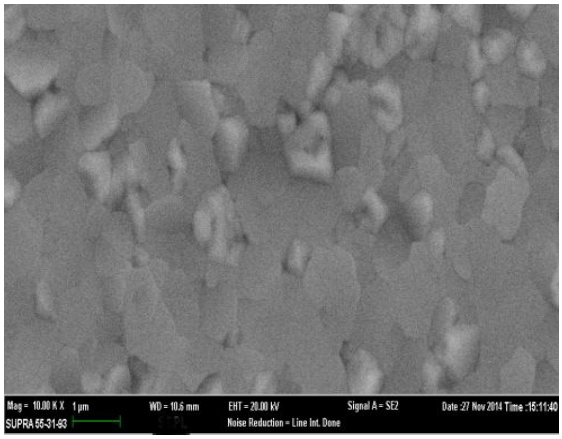


Fig: 17. SEM Micrographs of tin at 10X Magnification

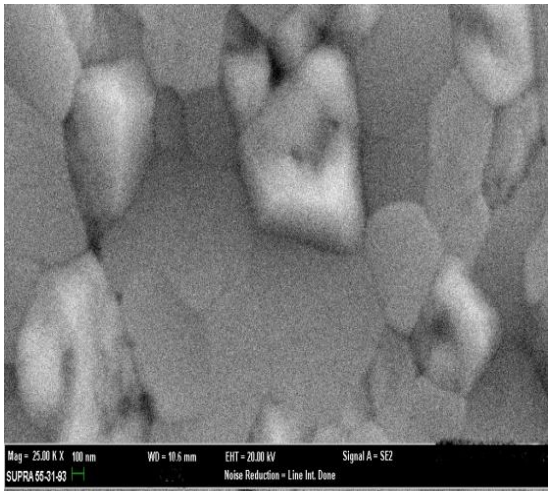


Fig: 18. SEM Micrographs of Tin at 25X Magnification

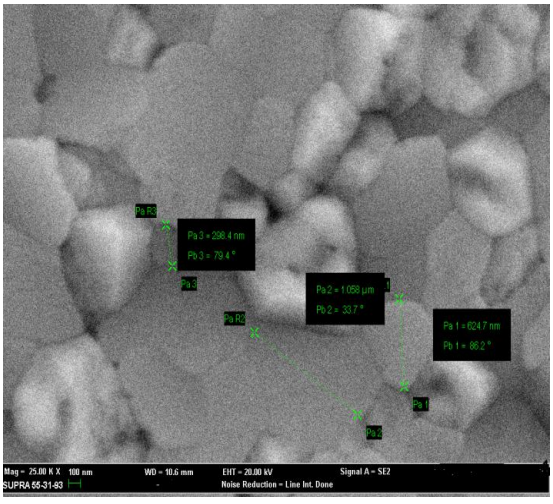


Fig: 19. SEM Micrographs of tin at 25X Magnification

Figure 20, 21 shows the SEM micrographs of Indium deposition on silicon substrate at 5KX,10KX magnification,20KV EHT, working distance 8.5mm for study the microsture and grain size and shows the grain sizes 10-200nm with coarse dispersion on the surface.

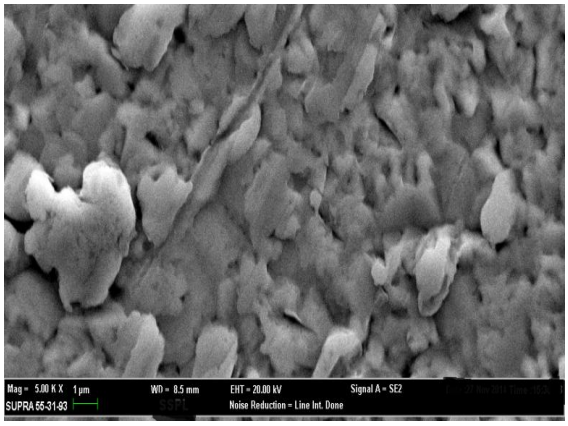


Fig: 20. SEM Micrographs of Indium-tin at 5X Magnificaion

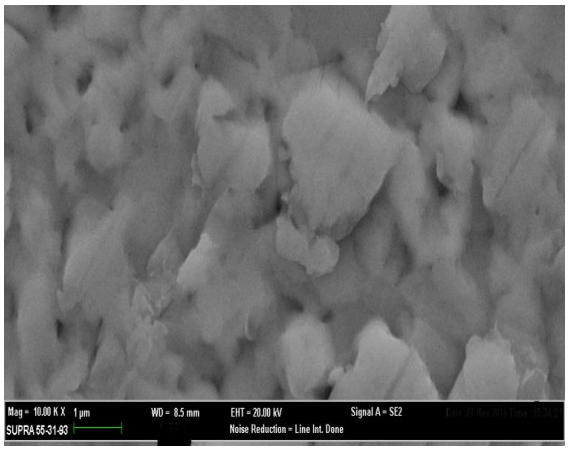


Fig: 21. SEM Micrographs of Indium-tin at 10X Magnificaion

Figure 22 shows the XRD analysis by using Kα (wavelength 1.541874) of tin on silicon which shows that it is polycrystalline structure with multiple peaks at 2θ angle

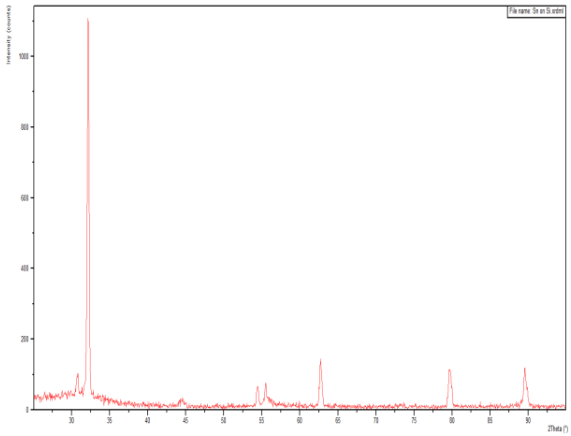


Fig: 22. XRD analysis of tin on silicon

Similarly, the XRD analysis of indium on silicon is shown in figure 23 which shows different peaks of intensity at different 2 angle revealing the polycrystalline structure.

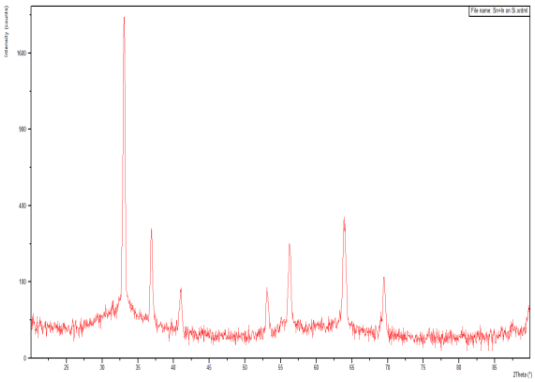


Fig: 23. XRD Analysis of Indium on Silicon

6. Conclusion

For Optimized Methods for Fast and Accurate Simulations, gases at low pressures cannot be modeled using conventional computational fluid dynamics tools. That is due to the fact that kinetic effects become important as the mean free path of the gas molecules becomes comparable to the length scale of the flow. Coatings of tin & indium having good uniformity & appropriate thickness have been produced on stationary, non-rotated, flat substrates using a vacuum thermal evaporation technique that incorporates the use of a resistively heated tungsten boat. The thickness uniformity around the substrate was a sensitive function of the mean free path as well as evaporant incidence angle. Also to increase the rate of evaporation we have to raise the temperature of the source. Thus this work presents the thermal evaporation of tin and indium on silicon with thickness of polycrystalline tin 164 nm and indium 182 nm. Also microstructure study at different magnifications show polycrystalline structure for tin and indium which can be transformed to single crystal and uniform grain size across grain boundaries by optimizing deposition rate and conditions. SEM micrographs of tin and indium at different magnifications shows the 100nm to 1microns grain size along the grain boundries. Similary, XRD analysis with Kα shows the peaks of intensity at different 2θ angles for different orientations of planes with polycrystalline structure.

References

[1] Ali Moarrefzadeh et.al., Simulation and Modeling of Physical Vapor Deposition (PVD) Process, Wseas Transactions on Applied and Theoretical Mechanics, 7(2), 2012
[2] G. H. Gilmer et.al., Thin film deposition: fundamentals and modeling, Journal of Computational Materials Science 12, 1998, 354-380
[3] C. R. Zamarreñoa et al. Optical Fiber Refractometers based on Indium Tin Oxide Coatings with Response in the Visible Spectral Region
[4] Souad Laghrib et al. , Tin oxide thin layers obtained by vacuum evaporation of tin and annealing under oxygen flow, Vacuum 82 (2008) 782-788
[5] Sea-Way Jan, Si-Chen Lee, Preparation and Characterization of Indium-Tin-Oxide Deposited by Direct Thermal Evaporation of Metal Indium and Tin, J. Electrochemistry. Soc.: Solid State Science and Technology, 134(8), 1987
[6] D. Bruce Buchholz et.al, Differences between amorphous indium oxide thin films; Progress in Natural Science: Materials International, 23(5), 2013, 475-480
[7] Alexandru C. Fechete et al., Growth of Indium Oxide Nanostructures by Thermal Evaporation, Sensor Technology lab., RMIT university, Australia, 1-4244-0453-3/06/$20.00@2006 IEEE
[8] C. A. Pan, T. P. Ma, High quality transparent conductive indium oxide films prepared by thermal evaporation, Appl. Phys. Lett., American Institute of Physics, 37(2), 1980
[9] B. Maniscalco, P. M. Kaminski, J. M. Walls, Thin film thickness measurement using Scanning White Light Interferometry, 550, 2014, 10-16
[10] M. Trzcinski et al, Alloy formation of ultrathin indium and silver layers on W(100): A photoemission study, Material Science 589 (2005) 192 -200
[11] Authur K. Burak Ucer, Vacuum evaporation, www.users.wfu.edu

Using Application of Dezert- Smarandache Theory on a new framework for protecting MANET

Ekata Gupta
MCA Department ,GNIM, New Delhi and
Research Scholar, Mewar University

Dr.S.K.Saxena
CSE Department
Delhi Technological University(Formerly Delhi College of Engineering)

Abstract - The article presents a concept of Dezert-Smarandache theory application for enhancing security in tactical mobile ad- hoc network. Tactical MANET, due to its specification, requires collection and processing of information from different sources of diverse security and trust metrics. The authors specify the needs for building a node's situational awareness and identify data sources used for calculations of trust metrics. They provide some examples of related works and present their own conception of Dezert-Smarandache theory applicability for trust evaluation in mobile hostile environment.

Keywords- situational awareness, trust, inference methods, tactical MANET, security, Dezert-Smarandache theory

I. INTRODUCTION

The history of mobile ad hoc networks came back to Packet Radio Networking project in 1972 that was done by DARPA The mobile ad-hoc networks are collections of independent nodes that can communicate via radio channels. These networks are often developed in conditions of limited or total lack of access to fixed infrastructure.

Mobile Ad hoc Network is an autonomous system of mobile devices that do not relay on any fixed infrastructure (see Fig.1). On one hand, lack of any fixed infrastructure makes these networks suitable for some critical applications such as emergency operations, disaster relief efforts, business indoor applications, civilian outdoor applications, military networks and so on . On the other hand, providing security for MANETs would be one of the most significant challenging issues.

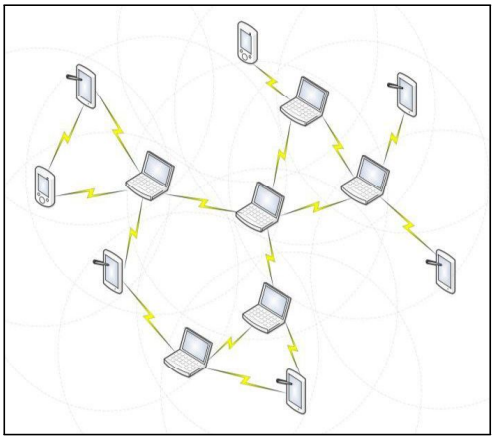


Figure 1. A sample mobile ad-hoc network structure

ISSN: 2319 – 1058

Security ensuring is particularly difficult for a tactical ad- hoc network, due to the necessity of dealing with a hostile environment, strict capacity constraints, the requirements for services, very rapid changes of network topology and dynamically forming groups of common interests, which cannot be pre-defined by trust relationships [1]. These networks are characterized by simple capability of adding new nodes, which may be of diverse nature, such as the allies, neutral or hostile nodes.

One method of ensuring the security is user authentication. Only the authorized nodes and those verified as allies can have access to the network. However, during the mission, a node can be taken over by the enemy, or change the nature of its behaviour - behaving to the detriment of the mission.

Due to the lack of a central management system it is needed for nodes to cooperate. Each of them is in fact a router ensuring cooperation between subnets and nodes located at a distance greater than the radio range.

Restrictions on ad-hoc networks contribute to the need of using other means than in wired networks to satisfy the safety requirements. In addition to authorization and authentication mechanisms, it is necessary for a node to have the knowledge on the behaviour of other nodes in the network, determining safety routes for data transfer and knowledge concerning the reaction manners in certain situations. The situational awareness building method will be complement of standard security mechanisms in mobile ad-hoc networks.

II. MOBILE AD HOC NETWORK CHARACTERISTICS

(a) Self-configured: Based on this feature, network functions (e.g. routing) must be done by the nodes, to keep the network usable.
(b) Dynamic topologies: Random movements of the nodes make the topology of the network unpredictable and highly dynamic.
(c) Resource constrained devices: Nodes are resource constrained in bandwidth, battery, memory, processing capability and computing power.
(d) Poor physical security: There is a possibility that mobile devices being captured or broken by adversaries.

III SECURITY ISSUES IN MOBILE AD HOC NETWORKS

Mentioned features above can lead to a set of underlying assumptions and performances and security concerns, which makes the design procedure more challenging. Due to the ability of communication over wireless channels and rapid growth of mobile devices, lack of widely accepted security solutions for this environment, made the research over security of MANETs still an active area .

Effective and proven security solutions for traditional networks are not always applicable to MANETs. In fact, attacks (e.g. identity/address spoofing, message tampering/ forgery/ replay) which may be easily detected and prevented in the traditional networks could be very challenging in MANETs. Mentioned issues in mobile ad hoc networks led to the difficulty of achieving security requirements .

Similarly to the traditional networks, MANETs need some security requirements which are described below :

(i) Data Confidentiality: keeping data secret from unauthorized entities
(ii) Data Integrity: preventing modification of data by unauthorized entities
(iii) Data Freshness: Keeping data in the correct order and up-to-date
(iv) Data Availability: Ensuring that data will be available on request
(v) Data& Identity Authentication: Verifying that the data or request is coming from a valid sender
(vi) Non-repudiation: Ensuring that a node cannot deny sending a message.

IV. NODE'S SITUATIONAL AWARENESS

A. Definitions

To identify opportunities of secure cooperation between nodes in ad-hoc networks, it is necessary to collect information about other nodes in the network. The ability to have accurate information about the surrounding reality and interpretation of the current situation in terms of the performed tasks is defined as a node's situational awareness.

The main product of the node's situational awareness mechanism is information on the node trust levels.

Trust is an interdisciplinary concept, characterized by a variety of definitions. It is understood as relying on the integrity, strength and ability of a person or thing. In the case of ad-hoc network it is translated as a set of relationships between people who use similar communication protocols [1]. These relations are defined based

ISSN: 2319 – 1058

on previous interactions of individuals. In [2], trust is treated as the degree of belief about the behaviour of other entities. Trust can also be understood as reputation, opinion, or the probability of correct behaviour [3].

In MANET, trust is the level of faith, which can be assigned by the node to its surroundings on the basis of observations and opinions coming from the other nodes in the network [4].

B. Security risk

Security risk and vulnerabilities assessment has many benefits and challenges associated with it. The security risk assessment should provide complete view of the existing security risk and help provide alternative solution and changes to the security measures and controls. We propose that an enhancement is needed, which will improve the security risk assessment
process, and that enhancement can be made to security risk management approach

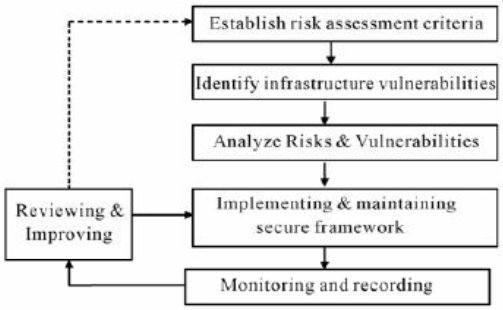


Figure 2: Security Risk Management framework

Secure exchange of information between nodes requires proper selection of the route of data transfer. Sending data via routes that are not safe may contribute to the leak or acquisition of data by unauthorized persons.

The dynamic process of creating a current situational view of node can be the basis for decisions on how to control traffic.

C. Data sources

Node's situational awareness in most cases is built based on direct interactions, indirect observations and recommendations.Trust determined by the node based on direct interaction and observation of behaviour of other nodes is called direct trust.

Trust determined on the basis of indirect observations and recommendations is called indirect trust. Recommendations shall be understood as opinions of other nodes on the node for which the level of confidence is being specified.

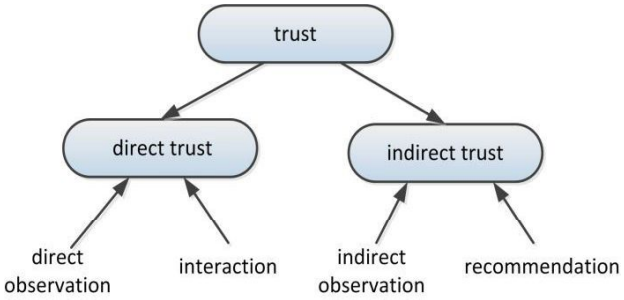


Figure 3. Direct and indirect trust

In many cases information from various sources may be incomplete, inconsistent or conflicting. This requires the selection of appropriate methods of inference, which would allow clear and accurate assessment of the current environment in which network node operates.

ISSN: 2319 – 1058

V. CONCEPT DESCRIPTION

Dynamic evaluation of the environment surrounding the node is possible by continuously monitoring the node behaviour, with the propsed frame work (Figure 2)

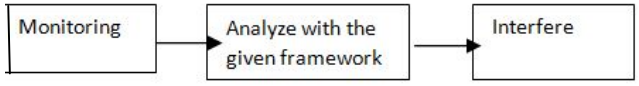


Figure 4. Nodes security evaluation process

In many cases, the knowledge acquired by a single node is insufficient to fully assess the current situation, therefore it must be able to exchange information about situational awareness built between nodes. Nodes can have different access to data about other nodes, so their passing information may be incomplete or uncertain. In the solutions described in section III, in most cases it is impossible to distinguish ignorance from uncertain knowledge, taking into account incomplete and conflicted knowledge derived from various sources.

The Dezert-Smarandache theory [12-14] allows combining information from multiple sources. It focuses on the problems of combining uncertain, conflicted and inaccurate information [15].

A. Events monitoring

Node assessment is made based on direct node observation and information from neighbouring nodes. Examples of observed events by which nodes can be evaluated are:

- provision of information - some of the nodes in ad-hoc networks are characterized by self-interested behaviour in order to deprive other nodes of the shares, for example by failing to forward packets for selfish node to the other nodes. Validation of packet transmission is possible through the analysis of incoming acknowledgments, when transmission of acknowledgments is enabled in the network or by tracking the packages sent by the monitoring node.
- compliance of safety rules - in tactical networks information may have different levels of sensitivity, for example: secret, confidential, non-confidential. Data on a certain level of sensitivity can be sent only to nodes that have access to information about a specific level or a higher level. Based on information collected on nodes access levels and data contained in the labels, it can be verified if a node observes the principles of safety, i.e. whether it has access only to data which is authorized and makes it available only to the authorized users.
- recommendation correctness - in the case when trust level is determined by recommendations from other nodes in the network, it is necessary to provide protection against “liar” nodes. A “liar” shall construe nodes, which transmit incorrect recommendations on other nodes, the objective of re-routing packet forwarding, intercepting or preventing delivery to the destination node.

The observed events can be evaluated as 0, 1 - using the classical theory of probability. However, in many cases, the observed behaviour provides some indication of both hypotheses, which would require omitting the evaluation of the event or a need to assign two assessments - which would misrepresent the two behaviours. Each behaviour is treated equally and the designated level of trust makes it impossible to identify the appropriate response to behaviour.

IV.CONCLUSION

Ensuring security in tactical MANET requires gathering and processing information about the node surrounding reality. Information from various sources, however, is often uncertain, incomplete and even conflicting. The method ensuring coverage of all of this information is Dezert- Smarandache theory, which allows representing of imprecise hypotheses. By applying the Dezert-Smarandache theory it is possible to identify specific and general hypotheses, which can combine data from different sources with access to information on the behaviour of nodes.

REFERENCES

[1] K. Seshadri Ramana, A.A. Chari, N. Kasiviswanth: “A Survey on Trust Management for Mobile Ad Hoc Networks”, International

ISSN: 2319 – 1058

Journal of Network Security & Its Applications (IJNSA), Vol. 2, No. 2, April2010.
[2] L. Capra: "Towards a Human Trust Model for Mobile Ad-hoc Networks", Dept. of Computer Science, University College London.
[3] Z. Han, K. J. R. Liu, Y. L. Sun, W. Yu: "A Trust Evaluation Framework in Distributed Networks: Vulnerability Analysis and Defence Against Attacks", INFOCOM 2006. 25th IEEE International Conference on Computer Communications, April 2006.
[4] J. Głowacka: "Procedures of building nodes' awareness for security in tactical ad-hoc networks, KKRRiT 2011, Poznań 2011, Telecommunication Review – Telecommunication News 2011 [CD], No. 6, pp. 405-408 (in Polish).
[5] Zhu Han, Yan Lindsay, K. J. Ray Liu, Wei Yu: "Information Theoretic Framework of Trust Modelling and Evaluation for Ad Hoc Networks",IEEE Journal on Selected Areas in Communications, Vol. 24, No. 2, February 2006.
[6] Jien Kato, Jie Li, Ruidong Li: "Future Trust Management Framework for Mobile Ad Hoc Networks", IEEE Communications Magazine, April 2008, pp. 108-114.
[7] C. Zouridaki, B. L. Mark, M. Hejmo, R. K. Thomas: "A Quantitative Trust Establishment Framework for Reliable Data Packet Delivery in MANETs", In SASN '05: Proceedings of the 3rd ACM workshop on Security of ad hoc and sensor networks, pp. 1–10, New York, NY, USA, 2005.
[8] V. Balakrishnan, V. Varadharajan, U. K. Tupakula, P. Lucs: "Trust and Recommendations in Mobile Ad hoc Networks", Third International Conference on Networking and Services, IEEE 2007.
[9] Junhai Luo, Xue Liu, Mingyu Fan,: "A trust model based on fuzzy recommendation for mobile ad-hoc networks", Computer Networks 53 (2009), pp. 2396–2407.
[10] Shafer G.: "A mathematical theory of evidence", Princeton U.P., Princeton, NJ, 1976.
[11] J. Konorski, R. Orlikowski: "DST-Based Detection of Noncooperative Forwarding Behavior of MANET and WSN Nodes", Proc. 2nd Joint IFIP WMNC., Gdansk, Poland, 2009.
[12] F. Smarandache, J. Dezert: "Advances and Applications of DSmT for Information Fusion", Vol 1, American Research Press Rehoboth, 2004.
[13] F. Smarandache, J. Dezert: "Advances and Applications of DSmT for Information Fusion", Vol. 2, American Research Press Rehoboth, 2006.
[14] F. Smarandache, J. Dezert: "Advances and Applications of DSmT for Information Fusion", Vol. 3, American Research Press Rehoboth, 2009.
[15] J. Głowacka, M. Amanowicz: „Situational awareness of a military MANET node – the basis" („Podstawy tworzenia świadomości sytuacyjnej węzła wojskowej sieci MANET"), Telecommunication Review – Telecommunication News 2012, No. 2-3, pp. 59-62 (in Polish).
[16] S. Bajpai, A. Sachdeva, J. Gupta, "Security Risk Assessment:Applying the Concept of Fuzzy Logic", *Journal of Hazardous Materials*, Vol. 173, No. 1-3, January 2010, pp.258-264.doi:10.1016/j.jhazmat.2009.08.078
[17] A. Veiga, J. Eloff, "A Framework and Assessment for Information Security Culture", *Computer and Security*,Vol. 29, No. 2, March 2010, pp. 196-207.doi:10.1016/j.cose.2009.09.002
[18] Zakaria I. Saleh, Heba Refai, Ahmad Mashhour "Proposed Framework for Security Risk Assessment" *Journal of Information Security*, 2011, 2, 85-90
[19] tameem eissa, shukor abd razak, md asri ngadi "*Enhancing MANET Security using Secret Public Keys"* 2009 International Conference on Future Networks
[20] Joanna Głowacka, Marek Amanowicz " Application of Dezert-Smarandache Theory for Tactical MANET Security Enhancement"

ISSN: 2319 – 1058

Analog Integr Circ Sig Process
DOI 10.1007/s10470-015-0546-7

MIXED SIGNAL LETTER

Voltage differencing transconductance amplifier based resistorless and electronically tunable wave active filter

Neeta Pandey[1] · Praveen Kumar[2] · Sajal K. Paul[3]

Received: 23 November 2014 / Revised: 8 April 2015 / Accepted: 18 April 2015
© Springer Science+Business Media New York 2015

Abstract This paper presents a systematic approach for the realization of resistorless wave active filter using voltage differencing transconductance amplifier. The proposed filter uses grounded capacitor and also possesses electronic tunability of cut-off frequency via bias current. The functionality of the proposed filter is verified for a 4th order low pass filter through PSPICE simulation using 0.18 μm TSMC CMOS technology parameter. Experimental results using commercially available IC LM13700/NS are also included which corroborate the theoretical and simulated results.

Keywords Voltage differencing transconductance amplifier · Wave active filter · Resistorless filter

1 Introduction

The high order passive RLC ladder filters are widely known for their low pass band sensitivity [1]. So, different methods such as operational simulation, element replacement and wave method have been developed for their active RC realization. In the first method [1], a signal flow graph (SFG) is constructed from the voltage–current relationships of prototype ladder. Both lossy and lossless integrators are used to implement SFG. The practical realization of lossless integrators is difficult because of non-ideality imposed by the used active and passive components. The complexity increases as the order of filter increases. In the second method, element replacement approach, the inductors of the passive prototype filters are replaced by impedance simulators designed using active elements such as voltage differencing transconductance amplifier (VDTA) [2, 3], current follower transconductance amplifier (CFTA) [4] and current backward transconductance amplifier (CBTA) [5]. The requirement of floating capacitor in the derived filter topology degrades the performance in high frequency application.

In wave active approach [6, 7], the third method, filter realization is based on the modeling of incident and reflected voltage wave. Each element of passive prototype filter is replaced by its wave equivalent. The advantages of this approach are—the derived filter structure is modular and it uses only lossy integrator. These features have led researchers to realize wave active filter structure using different active blocks. In particular, operational amplifier [6, 7], operational transconductance amplifier (OTA) [8], CFOA [9], DVCCTA [10], DVCCCTA [11], CCDDCCTA [12], OTRA [13], VDTA [14] have been explored to design higher order filters based on this approach.

In this paper, voltage differencing transconductance amplifier (VDTA) based realization of wave active filter is developed. The series inductor connection is used as an elementary building block and its wave equivalent is designed first. Wave equivalents of other possible ladder

✉ Neeta Pandey
n66pandey@rediffmail.com

Praveen Kumar
praveen.dtu@gmail.com

Sajal K. Paul
sajalkpaul@rediffmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University, Bawana Road, Delhi 110042, India

[2] Department of Electronics and Communication Engineering, IMS Engineering College, Ghaziabad, India

[3] Department of Electronics Engineering, Indian School of Mines, Dhanbad 826004, Jharkhand, India

Published online: 05 May 2015

elements e.g. series/parallel capacitor, parallel inductor etc. can be obtained by appropriately configuring the connections and adding an inverting buffer, if necessary. The VDTA provides resistorless filter realization due to presence of two internal transconductances. In addition, the filter parameters can be tuned electronically through bias currents of VDTA.The proposed filter structure has following attractive feature as compare to the structure of [14]:

- lesser active blocks per wave equivalent,
- resistorless realization,
- electronic tunability of cut-off frequency,
- does not employ inverters for cascading wave equivalents.

As an example a fourth order Butterworth low pass filter has been design by cascading VDTA based wave equivalents. The functionality is verified through PSPICE simulation using 0.18 μm TSMC technology parameter.Experimental results using commercially available IC LM13700/NS are also included.

2 VDTA

Voltage differencing transconductance amplifier is a current mode building block, which can be compared with current differencing transconductance amplifier (CDTA) [15]. In VDTA [16], differential input voltage, rather than current as in CDTA, is transferred to current at Z terminal by first transconductance gain and the voltage drop at the terminal Z is transferred to current at the terminal X by second transconductance gain. Both transconductances are electronically controllable by external bias currents. This new element exhibits two different values of transconductances so that several applications such as biquad filters, oscillator, inductance simulator and FDNR (frequency dependent negative resistor) simulator can be realized with a single active block employing one or two capacitors [2, 3, 16–26].

The Schematic symbol of VDTA is shown in Fig. 1 and the port relationship of the VDTA may be characterized by the following matrix.

$$\begin{bmatrix} I_Z \\ I_{X+} \\ I_{X-} \end{bmatrix} = \begin{bmatrix} g_{m1} & -g_{m1} & 0 \\ 0 & 0 & g_{m2} \\ 0 & 0 & -g_{m2} \end{bmatrix} \begin{bmatrix} V_P \\ V_N \\ V_Z \end{bmatrix} \quad (1)$$

In Eq. (1) g_{m1} and g_{m2} are the transconductances of input and second stage respectively. The differential input voltage from the terminals P and N is transformed into current at the terminal Z with first transconductance g_{m1}. The voltage drop at the terminal Z is converted into output

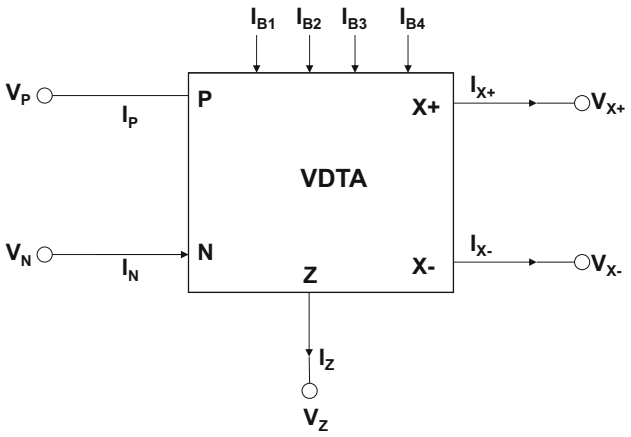


Fig. 1 Schematic symbol of VDTA

currents at the terminals X+ and X− with second transconductance g_{m2}.

The CMOS realization of VDTA is shown in Fig. 2. The circuit uses two Arbel–Goldminz transconductances. First and second transconductances of VDTA element in the circuit can be expressed as

$$g_{m1} = \frac{g_1 + g_5}{2} \quad \text{or} \quad g_{m1} = \frac{g_2 + g_6}{2} \quad \text{and} \quad (2)$$

$$g_{m2} = \frac{g_3 + g_7}{2} \quad \text{or} \quad g_{m2} = \frac{g_4 + g_8}{2}, \quad (3)$$

where g_i is the transconductance of ith transistor and can be expressed as

$$g_i = \sqrt{I_{Bi} C_{ox} \mu_i \left(\frac{W}{L}\right)_i}. \quad (4)$$

Here, I_{Bi} is the dc bias current of ith transistor, μ_i is the mobility of the carrier, C_{ox} is the effective gate oxide capacitance per unit area, W and L are the width and channel length of ith transistor.

3 Wave active filter

3.1 Realization of basic wave equivalents using VDTA

The resistively terminated LC ladder filter using wave method is based on representing two-port network in terms of incident and reflected waves. The scattering matrix description is used for this purpose which is subsequently employed for deriving corresponding wave equivalents. For a two port network of Fig. 3, the incidentand reflected voltage waves (A_j and B_j) are related by the following relations The resistively terminated LC ladder filter using wave method is based on representing two-port network in terms of incident and reflected waves. The scattering matrix

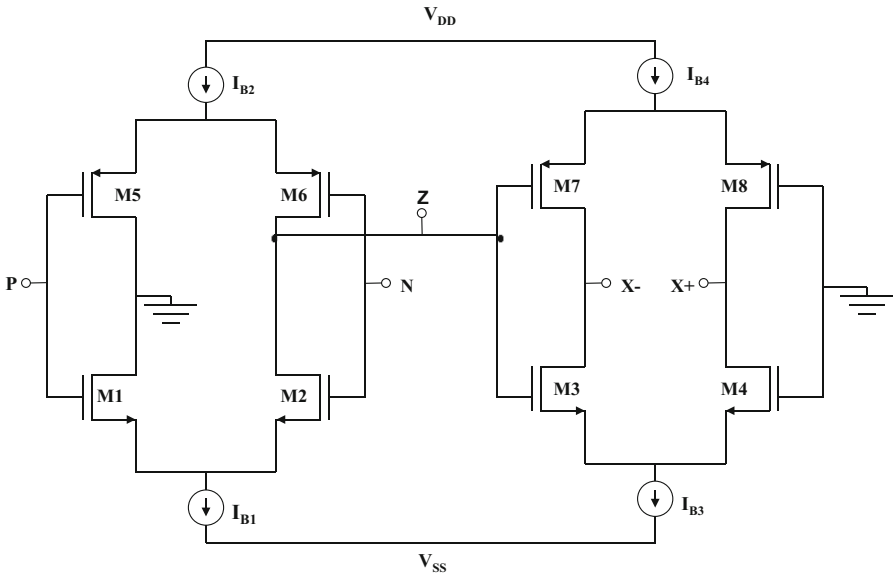


Fig. 2 CMOS based internal structure of VDTA

description is used for this purpose which is subsequently employed for deriving corresponding wave equivalents. For a two port network of Fig. 3, the incident and reflected voltage waves (A_j and B_j) are related by the following relations

$$A_j = V_j + I_j R_j, \quad B_j = V_j - I_j R_j \tag{5}$$

where j = 1, 2 and R_j is the resistance of port j.

In scattering matrix notation, (5) can be expressed in terms of scattering parameter(S) as

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = S \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \tag{6}$$

The basic element for developing wave active filter is a series inductor L. The scattering parameters for a series inductor are defined as

$$S = \frac{1}{1+s\tau} \begin{bmatrix} s\tau & 1 \\ 1 & s\tau \end{bmatrix} \tag{7}$$

The relationship between incident wave A_j and the reflected wave B_j of a series inductor may be obtained using (6) and (7) as

$$B_1 = A_1 - \frac{1}{1+s\tau}(A_1 - A_2) \tag{8}$$

$$B_2 = A_2 + \frac{1}{1+s\tau}(A_1 - A_2) \tag{9}$$

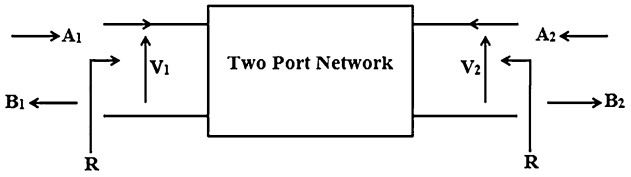


Fig. 3 Two port network with wave variables

where the time constant τ in terms of desired inductance L and port resistance R is expressed as

$$\tau = L/2R \tag{10}$$

The wave equivalent of series inductor can be realized by active implementation of (8) and (9). It requires three operations—differential lossy integration, subtraction and summation which can easily be realized by VDTA as shown in Fig. 4. The function of each VDTA of Fig. 4 is illustrated as follows by assuming $g_{m1} = g_{m2} = g_m$ for all the three VDTAs.

The first VDTA performs differential lossy integration and provides output voltage as

$$V_{01} = \frac{1}{1+s\tau}(A_1 - A_2), \tag{11}$$

$$\text{where} \quad \tau = \frac{C_V}{g_m} \tag{12}$$

Equating (10) and (12), the value of C_V may be computed as

$$C_V = \frac{g_m^2 L}{2} (\text{Assuming } Rg_m = 1) \tag{13}$$

The subtraction operation needed for active realization of (8) is obtained by second VDTA which provides output voltages as

$$B_1 = A_1 - V_{o1} \tag{14}$$

It may be noted that the third VDTA does not use $V_{01}$for summing operation. Instead, the current output from differential lossy integrator available at X+ port of first VDTA is connected to output terminal of third VDTA which performs operation equivalent to

$$B_2 = A_2 + V_{o1} \tag{15}$$

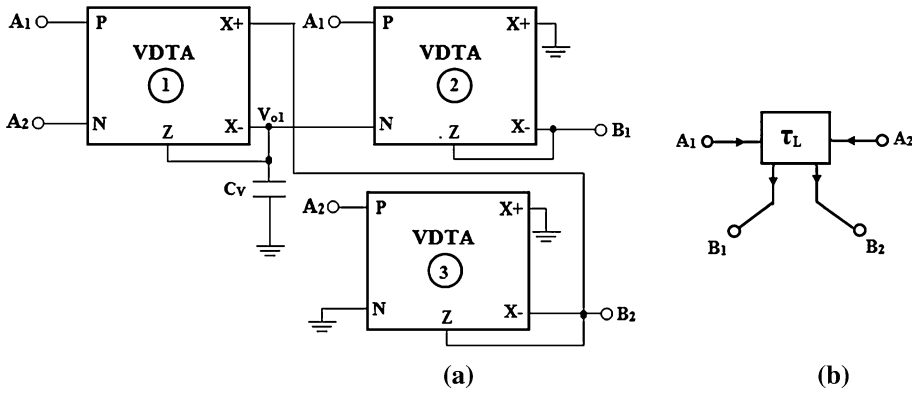
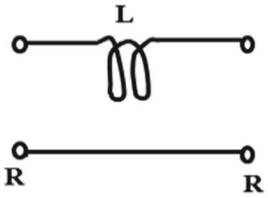
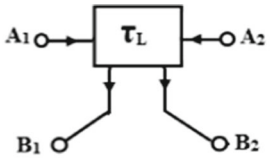
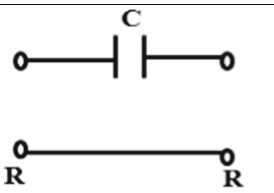
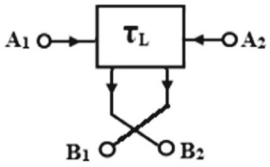
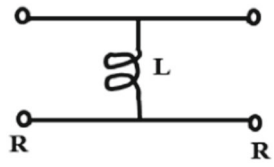
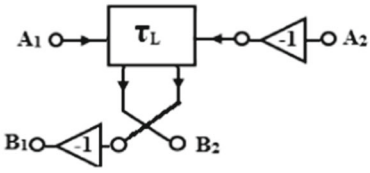
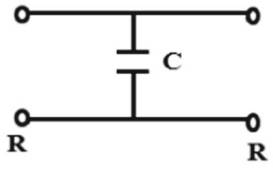
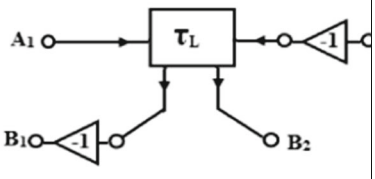


Fig. 4 **a** Complete schematic of VDTA based wave equivalent of series inductor. **b** Its symbolic representation

Table 1 Wave equivalent of elementary two ports consisting of single elements in series and shunt branch

Elementary two port network	Port connection	Realized time constant; required capacitor value (C_V) to implement corresponding VDTA based wave equivalent
		$\tau = \frac{L}{2R}$; $C_V = \frac{g_m{}^2 L}{2}$
		$\tau = 2RC$; $C_V = 2C$
		$\tau = \frac{2L}{R}$; $C_V = 2g_m{}^2 L$
		$\tau = \frac{RC}{2}$; $C_V = \frac{C}{2}$

The symbolic representation of series inductor wave equivalent is shown in Fig. 4(b).

3.2 Realization of other passive component

The structure shown in Fig. 4 can be used as the basic building block for deriving the wave equivalent of other reactive elements. The wave equivalent for series and shunt inductor and capacitor are given in Table 1 which can be obtained by swapping outputs and signal inversion. The VDTA circuit for signal inversion is shown in Fig. 5. The design of wave active filter starts with the selection of prototype filter based on specifications. The individual inductors or capacitors are replaced by their wave equivalents given in Table 1 [6, 7]. The complete filter schematic is then obtained by simply cascading the wave equivalents.

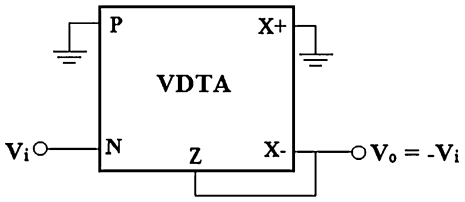


Fig. 5 VDTA based inverter circuit

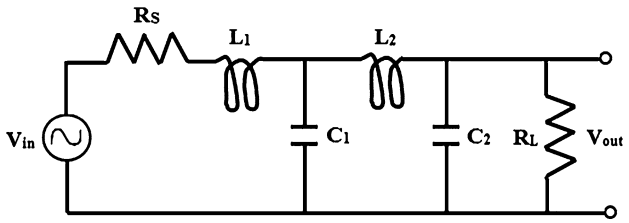


Fig. 6 A 4th order Butterworth filter

3.3 Design example

To illustrate the design procedure a Fourth order filter of Fig. 6 is considered. To derive the wave equivalent topology of Fig. 6 the series inductor and shunt capacitor may be replaced by wave equivalent as given in Table 1. The resulted wave equivalent topology is shown in Fig. 7 which uses four τ_L blocks and four inverters. A careful examination reveals that the inverters may easily be avoided by modifying the wave equivalent. If outputs of τ_L are taken as—B_1 and B_2 and abbreviating the resulting block as τ_{L_inv}, the topology of Fig. 7 modifies to Fig. 8 which makes use of only one τ_L and three τ_{L_inv} blocks. The VDTA based realization of τ_{L_inv} block is depicted in Fig. 9 where the connections remain same as the one used in τ_L except in 2nd VDTA where inputs are interchanged. Thus, the introduction of τ_{L_inv} saves four VDTAs for this example and subsequently results in lesser power consumption.

4 Non ideal analysis

In practice, (8) and (9) modify due to non idealities which are classified as parasites and tracking errors. The parasites in form of shunt output impedances (R//C) are present at terminals P, N, Z, X+ and X− and are denoted respectively as (R_P//C_P) (R_N//C_N) (R_Z//C_Z) (R_{X+}//C_{X+}) (R_{X-}//C_{X-}). The effect of the parasites is highly dependent on the topology. A close inspection of the circuit of Fig. 4(a) shows that the parasitic capacitances present at Z and X- terminals of 1st VDTA can be easily accommodated in external capacitance C_V. Moreover, parasites at terminals connected directly to inputs $A_1$and A_2 will also not cause any change. Considering these parasites, the

$$B_1 = \frac{A_1 - V_{o1}}{\left\{1+\frac{1}{g_m}\left(\frac{1}{R_Z}+\frac{1}{R_{X-}}\right)\right\}+\frac{s(C_Z+C_{X-})}{g_m}} \tag{16}$$

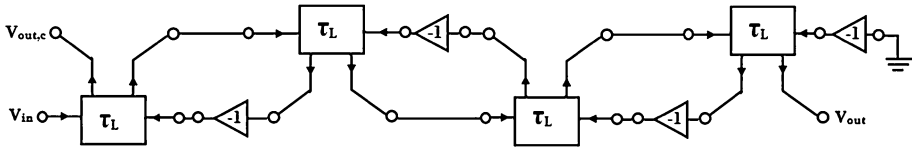


Fig. 7 Wave equivalent of prototype filter

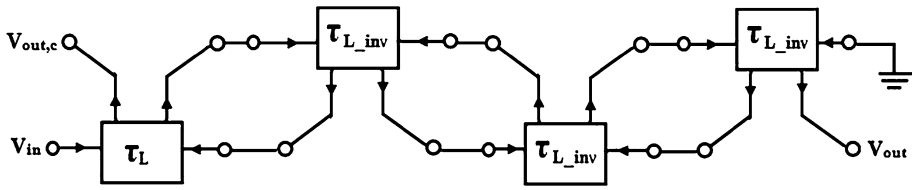


Fig. 8 New wave equivalent of prototype filter without use of inverter

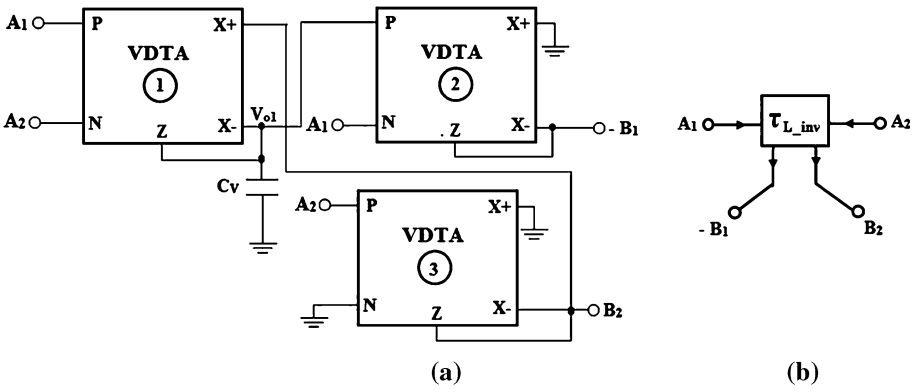


Fig. 9 **a** Complete schematic of VDTA based wave equivalent of series inductor with inverted B_1. **b** It's symbolic representation

and

$$B_2 = \frac{A_2 + V_{o1}}{\left\{1 + \frac{1}{g_m}\left(\frac{1}{R_Z} + \frac{1}{R_{X-}} + \frac{1}{R_{X+}}\right)\right\} + \frac{s\,(C_Z + C_{X-} + C_{X+})}{g_m}} \tag{17}$$

where V_{o1}, the output of first VDTA is given by

$$V_{o1} = \frac{1}{\left\{1 + \frac{1}{g_m}\left(\frac{1}{R_Z} + \frac{1}{R_{X-}} + \frac{1}{R_N}\right)\right\} + s\,\tau'}(A_1 - A_2) \tag{18}$$

$$\text{where} \quad \tau' = \frac{s(C_V + C_Z + C_{X-} + C_N)}{g_m} \tag{19}$$

The terms in curly braces in (16)–(18) may be approximated as unity as parasitic resistances at various ports have very high value and their reciprocal will be very small. It may be noted that the parasitic capacitances introduce poles at very high frequencies $\frac{g_m}{(C_Z + C_{X-})}$ and $\frac{g_m}{(C_Z + C_{X-} + C_{X+})}$, as the value of parasitic capacitance at z and x port are very small (0.15 pF [16]). The value of time constant is also modified which can be adjusted by bias current.

Second non-ideality comes due to error in transconductance transfer function from p and n ports to z ports and z ports to x+ and x− port [17]. The modified port relationship may be written as-

$$I_z = \beta_1 g_{m1}(V_P - V_N) \tag{20}$$

$$I_{X+} = \beta_2 g_{m2} V_Z \tag{21}$$

$$I_{X-} = -\beta_2 g_{m2} V_Z \tag{22}$$

β_1 and β_2 are the tracking errors for the first and second stages of the VDTA.

Considering these tracking errors, the modified equation for B_1 and B_2 are-

$$B_1 = \frac{\beta_1}{\beta_2}(A_1 - V_{o1}) \tag{23}$$

$$B_2 = \frac{\beta_1}{\beta_2} A_2 + V_{o1} \tag{24}$$

$$\text{where} \quad V_{o1} = \frac{\frac{\beta_1}{\beta_2}(A_1 - A_2)}{1 + \frac{sC_V}{g_m \beta_2}} \tag{25}$$

Now the values of C_V can be calculated from the equation

$$C_V = \frac{\beta_2 g_m^2 L}{2} \tag{26}$$

The effect of tracking error in (25) and (26) can be minimized by properly adjusting the VDTA's transconductance value with bias currents

5 Comparison

The proposed VDTA based wave equivalents are compared with the existing ones on the basis of active and passive elements used; and possibility of electronic tunabilty. The findings are summarized in Table 2. It may be noted that

- Grounded capacitor is used in all the structures [8–14]
- Additional active elements are needed to realize shunt wave equivalents [9–11, 14].
- Current duplicators are needed in [8] and passive or MOS based resistors are required in [13]
- Electronic tuning is not possible [9, 10, 14]

The wave equivalents of Ref [12] and the proposed one use only three active elements and are electronically tunable. The proposed wave filter is advantageous as internal structure of VDTA [16] is simpler than CCDDCCTA [12].

Table 2 Comparison of Wave Equivalents

Ref.	Active block	Wave equivalents	Active component used	Passive component used		Electronic tunabilty
				C	R	
[8]	OTA	Series inductor	2 OTA, 3 Duplicator	1 (G)	–	Yes
		Series capacitor	2 OTA, 3 Duplicator	1 (G)	–	
		Shunt inductor	2 OTA, 3 Duplicator	1 (G)	–	
		Shunt capacitor	2 OTA, 3 Duplicator	1 (G)	–	
[9]	CFOA	Series inductor	3 CFOA	1 (G)	4 (G) + 4 (F)	No
		Series capacitor	3 CFOA	1 (G)	4 (G) + 4 (F)	
		Shunt inductor	5 CFOA	1 (G)	4 (G) + 4 (F)	
		Shunt capacitor	5 CFOA	1 (G)	4 (G) + 4 (F)	
[10]	DVCCTA	Series inductor	4 DVCCTA	1 (G)	4 (G)	No
		Series capacitor	4 DVCCTA	1 (G)	4 (G)	
		Shunt inductor	6 DVCCTA	1 (G)	4 (G)	
		Shunt capacitor	6 DVCCTA	1 (G)	4 (G)	
[11]	DVCCCTA	Series inductor	4 DVCCCTA	1(G)	–	Yes
		Series capacitor	4 DVCCCTA	1 (G)	–	
		Shunt inductor	6 DVCCCTA	1 (G)	–	
		Shunt capacitor	6 DVCCCTA	1 (G)	–	
[12]	CCDDCCTA	Series inductor	3 CCDDCCTA	1 (G)	–	Yes
		Series capacitor	3 CCDDCCTA	1 (G)	–	
		Shunt inductor	3 CCDDCCTA	1 (G)	–	
		Shunt capacitor	3 CCDDCCTA	1 (G)	–	
[13]	OTRA	Series inductor	3 OTRA, 18MOS	1 (G)	–	Yes
		Series capacitor	3 OTRA, 18MOS	1 (G)	–	
		Shunt inductor	5 OTRA, 18MOS	1 (G)	–	
		Shunt capacitor	5 OTRA, 18MOS	1 (G)	–	
[14]	VDTA	Series inductor	4 VDTA	1 (G)	8 (G)	No
		Series capacitor	4 VDTA	1 (G)	8 (G)	
		Shunt inductor	6 VDTA	1 (G)	8 (G)	
		Shunt capacitor	6 VDTA	1 (G)	8 (G)	
Proposed	VDTA	Series inductor	3 VDTA	1 (G)	–	Yes
		Series capacitor	3 VDTA	1 (G)	–	
		Shunt inductor	3 VDTA	1 (G)	–	
		Shunt capacitor	3 VDTA	1 (G)	–	

* G = Grounded, F = Floating

Table 3 Aspect ratio of transistors used in VDTA

Transistors	Aspect ratios (W(μm)/L(μm))
M_1–M_2	3.6/.36
M_3–M_4	3.1/.36
M_5–M_8	16.64/3.6

6 Simulation

The functionality of the VDTA based wave filter is demonstrated by taking a 4th order low pass filter of Fig. 6 as prototype. For maximally flat response, the normalized component values are $R_S = 1$, $L_1 = .7654$, $L_2 = 1.8485$, $C_1 = 1.8485$, $C_2 = .7654$ and $R_L = 1$. The wave equivalent topology without inverter is shown in Fig. 8. The topology of Fig. 8 is simulated using VDTA based wave equivalent shown in Figs. 4 and 9, as discussed in Sect. 3 using 0.18 μm TSMC CMOS technology parameter. The aspect ratio of various transistors is given in Table 3. Supply voltages are taken as $V_{DD} = -V_{SS} = 0.9$ V and bias currents are taken as $I_{B1} = I_{B2} = I_{B3} = I_{B4} = 150$ μA. For cutoff frequency $f_o = 10$ MHz, the capacitor values for wave equivalent of series inductors (L_1, L_2) and shunt capacitor (C_1, C_2) are calculated to be 3.8, 9.2 pF and 9.2, 3.8 pF respectively. The low pass response (V_{out}) and

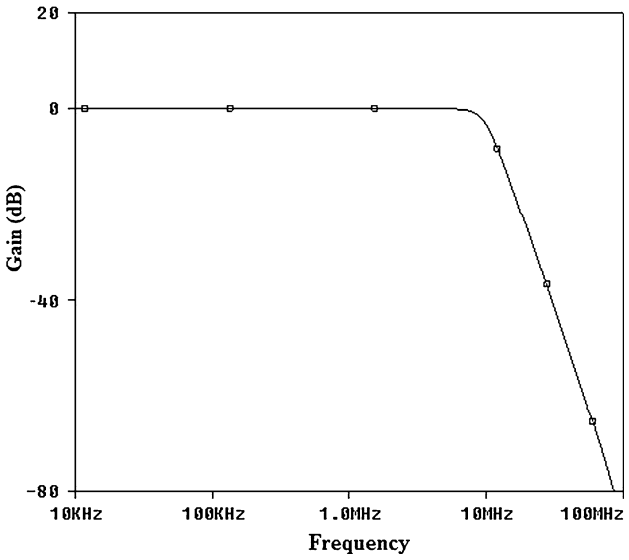


Fig. 10 Frequency response of 4th order low pass filter

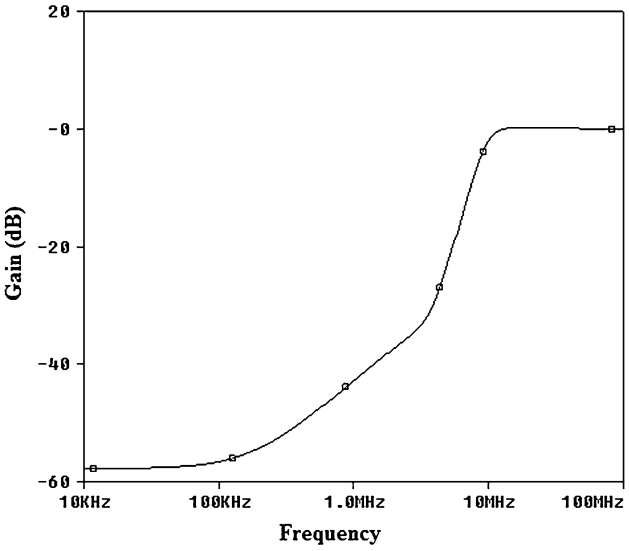


Fig. 11 Frequency response of complementry high pass filter

and its complementary high pass response ($V_{out.c}$) are shown in Figs. 10 and 11 respectively. Figure 12 shows the tunability of filter response through simulation by varying bias current from 20 to 250 μA. The power dissipation of VDTA based filter is 6.48 mW. To study the time domain behavior, input signals comprised of two frequencies of 5 and 30 MHz and amplitude of 50 mV is applied. The transient response and its spectrum for input and output are shown respectively in Figs. 13 and 14, which clearly shows that the 30 MHz signal is significantly attenuated. The proposed circuit is also tested to judge the level of harmonic distortion at the output of the signal. The %THD is within acceptable limit of 1.5 % up to about 800 mV p-p as shown in Fig. 15.

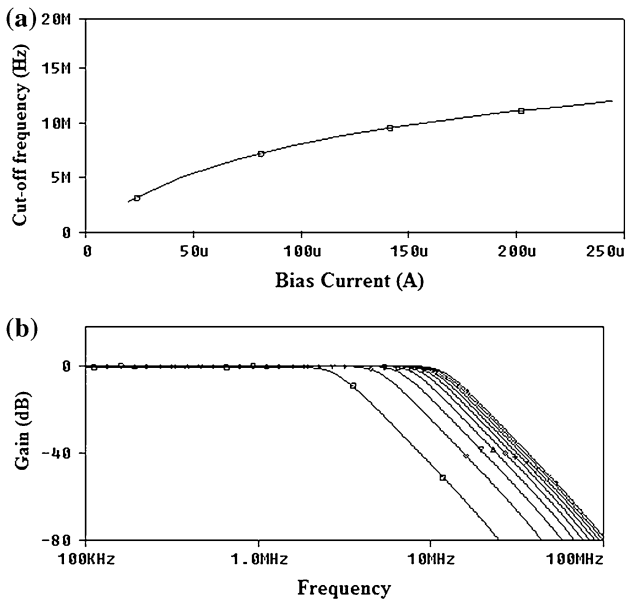


Fig. 12 Demonstration of electronic tunability **a** Cut-off tunability with bias current **b** Frequency response for different bias currents

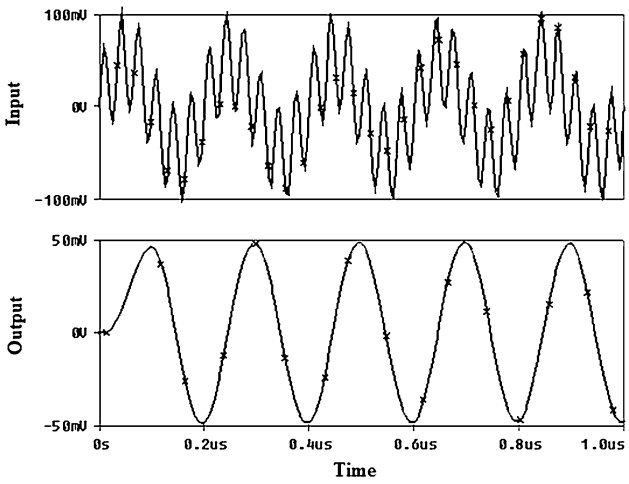


Fig. 13 Transient response of input and output signal for proposed filter

The performance of the proposed circuit is compared with the previous design of wave active structures [9–13] in terms of power consumption, %THD, output noise and electronic tunability. It may be noted from Table 4 that the proposed design shows good %THD result with lowest power consumption and lowest output noise voltage. It may be noted that the relevant data for structures of [6–8, 14] is not available in the literature.

The filter topology of [14] is also simulated for the same cut-off frequency of 10 MHz. The power dissipation for this circuit is 35.1 mW, output noise is 2.396×10^{-8} V/Hz$^{1/2}$ and %THD is within 1 % up to 300 mV p–p. These result shows that the proposed structure is better than the filter topology of [14].

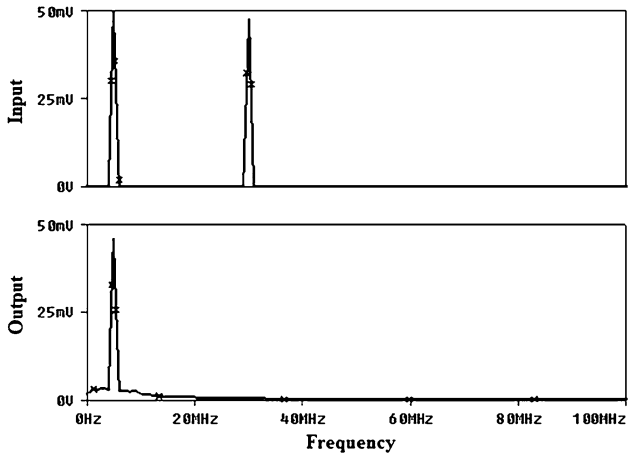


Fig. 14 Spectrum of input and output signal for proposed filter

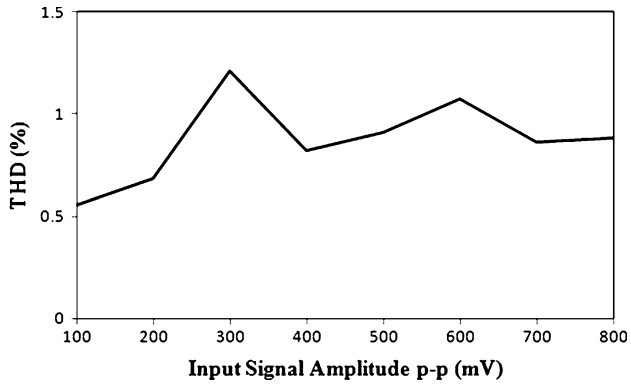


Fig. 15 Variation of %THD with input signal amplitude

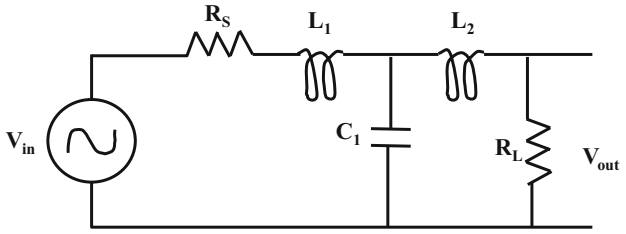


Fig. 16 abcd

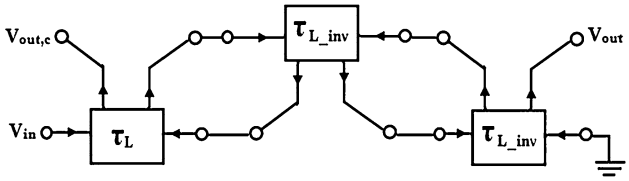


Fig. 17 Wave equivalent of 3rd order prototype filter

7 Experimental verification

The functionality of the proposed wave filter is experimentally demonstrated with a 3rd order low pass filter prototype of Fig. 16. For maximally flat response, the normalized component values are $R_S = 1, L_1 = 1, L_2 = 1, C_1 = 2$, and $R_L = 1$. The wave equivalent topology of Fig. 17 is obtained with the procedure described in Sect. 3. The wave equivalent topology of Fig. 17 is obtained with the procedure described in Sect. 3. The circuit is breadboarded for experimental testing wherein the functionality of the VDTA is implemented using commercially available IC LM13700/NS as shown in schematic of Fig. 18(a). The

Table 4 Comparison of proposed circuit with previous circuits

Ref.	Active block and technology used	Filter structure	Electronic tunabilty	Output noise voltage ($V/HZ^{1/2}$)	Power consumption (mW)	% THD
[9]	CFOA (Commercially available IC AD844 with ± 5 V power supply)	3rd order elliptic Low pass	No	N/A	N/A	1 % for 1 V pp signal
[10]	DVCCTA (CMOS technology 0.25 µm, power supply ± 1.25 V)	4th order Butterworth Low pass	No	9.48×10^{-8}	24.4	15 % up to 400mv pp signal
[11]	DVCCCTA (CMOS technology 0.25 µm, power supply ± 1.25 V)	4th order Butterworth Low pass	Yes	8.36×10^{-8}	59.2	5 % up to 225 mV pp signal
[12]	CCDDCCTA (CMOS technology 0.25 µm, power supply ± 1.25 V)	3rd order Butterworth low pass	Yes	6.73×10^{-8}	13.7	6 % up to 300 mV pp signal
[13]	OTRA (CMOS technology 0.5 µm, power supply ± 1.5 V)	3rd order Butterworth low pass	Yes	7.26×10^{-8}	10.7	Less than 5 % for 125 mV pp signal
Proposed	VDTA (CMOS technology 0.18 µm, power supply ± 0.9 V)	4th order Butterworth low pass	Yes	1.65×10^{-8}	6.48	Less than 1.5 % for 800 mV p–p signal

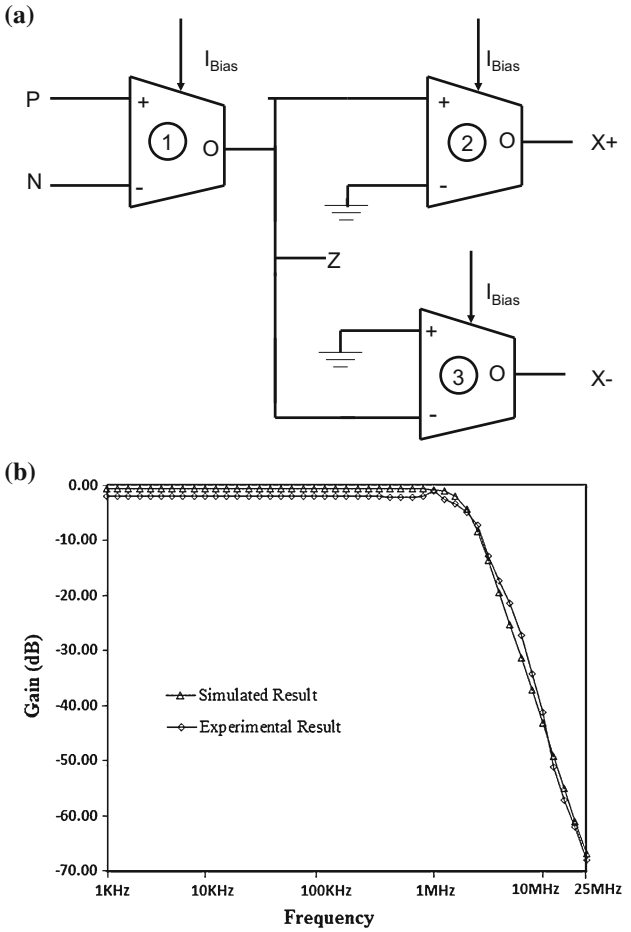


Fig. 18 **a** VDTA implementation using OTA. **b** Simulated and observed low pass response

Supply voltages of ±15 V are used. The bias current of 1.3 mA is chosen to set the value of g_m at 24.891 mA/V. The used capacitor values for wave equivalent of series inductors (L_1, L_2) and shunt capacitor (C_1) are 0.1, 0.1 and 0.2 nF, respectively. The simulation are also carried out using OTA IC LM 13700/NS. The simulated and observed low pass response (V_{out}) is shown in Fig. 18(b) and is in close agreement. The experimental value of cut-off frequency is observed to be 2.1 MHz as against the theoretical value of 1.98 MHz. The slight deviation in experimental results may be attributed to component tolerance of ±10 % as used in experiment.

8 Conclusion

A design technique for obtaining VDTA based wave active filter is presented. A VDTA based wave equivalent of series inductor is proposed as it is basic building block. The realization so obtained is then configured for other passive element realization by making suitable connections. The structure is resistorless, employs grounded capacitors, and possesses electronic tunability of cutoff frequency. The wave active method offer modularity and an ability to design higher order filter in short time. The theoretical proposition is verified for a fourth order low pass filter through PSPICE simulation using 0.18 μm TSMC CMOS technology parameters.There is close agreement between theoretical and experimental results.

References

1. Van Valkenburg, M. E., & Shaumann, R. (2001). *Design of analog filters*. Oxford (UK): Oxford University Press.
2. Prasad, D., & Bhaskar, D. R. (2012). Grounded and floating inductance simulation circuits using VDTAs. *Circuits and Systems, 3*(4), 342–347.
3. Tangsrirat, W., & Unhavanich, S. (2014). Voltage differencing transconductance amplifier-based floating simulators with a single grounded capacitor. *Indian Journal of Pure and Applied Physics, 52*(6), 423–428.
4. Li, Y. (2012). A series of new circuits based on CFTAs. *AEU - International Journal of Electronics and Communications, 66*(7), 587–592.
5. Ayten, U. E., Sagbas, M., Herencsar, N., Koton, J. (2012). Novel floating general element simulators using CBTA.*Radioengineering,* 21(1), 11-19.
6. Wupper, H., & Meerkotter, K. (1975). New active filter synthesis based on scattering parameters. *IIEEE Transcation on Circuit and System, 22*(7), 594–602.
7. Haritantis, I., Constantinides, A., & Deliyannis, T. (1976). Wave active filter. *IEEE Proceeding, 123*(7), 676–682.
8. Tingleff, J., & Toumazou, C. (1975). A 5th order lowpass currentmode wave active filter in CMOS technology. *Analog IntegratedCircuits and Signal Processing, 7*, 131–137.
9. Georgia, K., & Costas, P. (2010). Modular filter structures using CFOA. *Radio engineering, 19*(4), 662–666.
10. Pandey, N., & Kumar, P. (2011). Differential voltage current conveyor transconductance amplifier based wave active filter. *Journal of Electron Devices, 10*, 429–432.
11. Pandey, N., & Kumar, P. (2011). Realization of resistorless wave active filter using differential voltage current controlled conveyor transconductance amplifier. *Radioengineering, 20*(4), 911–916.
12. Pandey, N., Kumar, P., & Choudhary, J. (2011). Current controlled differential difference current conveyor transconductance amplifier and its application as wave active filter. *ISRN Electronics,*. doi:10.1155/2013/968749.
13. Bothra, M., Pandey, R., Pandey, N., & Paul, S. K. (2013). Operational trans-resistance amplifier based tunable wave active filter. *Radioengineering, 22*(1), 159–166.
14. Singh, H., Arora, K., & Prasad, D. (2014). VDTA—based wave active filter. *Circuit and System,*. doi:10.4236/cs.2014.55014.
15. Biolek, D. (2003). CDTA—Building block for current-mode analog signal processing. In: *Proceedings of ECCTD03,* Krakow (Poland), 397–400.
16. Yesil, A., Kacar, F., & Kuntman, H. (2011). New simple CMOS realization of voltage differencing transconductance amplifier and its RF filter application. *Radioengineering, 20*(3), 632–637.

17. Satansup, J., Pukkalanun, T., & Tangsrirat, W. (2013). Electronically tunable current mode universal filter using VDTAs and grounded capacitors. In: Proceedings of IMECS, Hong kong, (pp. 647–650).
18. Prasad, D., Srivastava, M., & Bhaskar, D. R. (2013). Electronically controllable fully-uncoupled explicit current-mode quadrature oscillator using VDTAs and grounded capacitors. *Circuits and Systems, 4*(2), 169–172.
19. Prasad, D., & Bhaskar, D. R. (2012). Electronically controllable explicit current output sinusoidal oscillator employing single VDTA. *ISRN Electronics,*. doi:10.5402/2012/382560.
20. Sotner, R., Jerabek, J., Herencsar, N., Petrzela, J., Vrba, K., & Kincl, Z. (2014). Linearly tunable quadrature oscillator derived from LC Colpitts structure using voltage differencing transconductance amplifier and adjustable current amplifier. *Analog Integrated Circuits and Signal Processing, 81*(1), 121–136.
21. Satansup, J., & Tangsrirat, W. (2014). Compact VDTA-based current-mode electronically tunable universal filters using grounded capacitors. *Microelectronics Journal, 45*(6), 613–618.
22. Prasad, D., Bhaskar, D. R., & Srivastava, M. (2013). Universal voltage-mode biquad filter using voltage differencing transconductance amplifier. *Indian Journal of Pure and Applied Physics, 51*(12), 864–868.
23. Yesil, A., & Kacar, F. (2013). Electronically tunable resistorless mixed mode biquad filters. *Radioengineering, 22*(4), 1016–1025.
24. Herencsar, N., Sotner, R., Koton, J., Misurec, J., & Vrba, K. (2013). New compact VM four-phase oscillator employing only single z-copy VDTA and all grounded passive elements. *Elektronika ir Elektrotechnika, 19*(10), 87–90.
25. Mekhum, W., & Jaikla, W. (2013). Three input single output voltage-mode multifunction filter with independent control of pole frequency and quality factor. *Advances in Electrical and Electronic Engineering, 11*(6), 494–500.
26. Uygur, A., & Kuntman, H. (2013). DTMOS-based 0.4V ultra low-voltage low-power VDTA design and its application to EEG data processing. *Radioengineering, 22*(2), 458–466.



Neeta Pandey received her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani and Ph.D. from Guru Gobind Singh Indraprastha University, Delhi. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is assistant professor in ECE Department, Delhi Technological University. A life member of ISTE, and member of IEEE, USA, she has published papers in international, national journals of repute and conferences. Her research interests are in analog and digital VLSI design.



Praveen Kumar is a part time research scholars in ISM, Dhanbad. He received his M.Tech from Delhi Technological University, Delhi and B.Tech from Uttar Pradesh Technical University, Lucknow. Currently he is working as an assistant professor in ECE department, IMS Engineering College, Ghaziabad. He also served as a lecturer in Hindustan College of Science and Technology, Faraha, Mathura for two year. His research area is analog VLSI Design.



Sajal K. Paul received his B.Tech., M.Tech. and Ph.D. in Radio Physics and Electronics from the Institute of Radio Physics and Electronics, University of Calcutta. He has served Webel Telecommunication Industries, Kolkata; Indira Gandhi National Open University (IGNOU), Kolkata; Advanced Training Institute for Electronics & Process Instrumentation (ATI-EPI), Hyderabad; North Eastern Regional Institute of Science & Technology (NERIST), Nirjuli and Delhi College of Engineering (DCE), Delhi in various capacities. He has served the Department of Electronics Engineering, Indian School of Mines, Dhanbad as head of the department and at present is a professor of the same department. His research interest includes microelectronic devices, electronic properties of semiconductor and bipolar and MOS analog integrated circuits. Dr. Paul has more than 70 research publications in international and national journals of repute and conferences.