CENTRAL LIBRARY

SCHOLARLY ARTICLES

A CURRENT AWARENESS BULLETIN OF ARTICLES BY FACULTY, STUDENTS AND ALUMNI

~ MAY 2013 ~

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 Fifth Issue of Current Awareness Bulletin for the year 2013, started by Delhi Technological University 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 DTU as a whole to the academia. It contains information resources available in the internet in the form of articles, reports, presentation published in international journals, websites, etc. by the faculty and students of Delhi Technological University in the field of science and technology. The publication of Faculty and Students, 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. To make the bulletin more comprehensive, the learned faculty and Students may provide their uncovered publication to the library either through email or in CD, etc.

This issue contains the information published during May 2013. The arrangement of the contents is alphabetical wise starting from A-Z. The Full text of the article, which is either subscribed by the University or available in the web, is provided in this Bulletin.

CONTENTS

11. Novel Applications of Functionally Graded Nano, Optoelectronic and Thermoelectric Materials ***Isha Bharti***, Nishu Gupta, and K. M. Gupta

12. Observational constraints on the kinematics and the fate of the Universe through linearly varying deceleration parameter laws Özgür Akarsu and Tekin Dereli, ***Suresh Kumar***, Lixin Xu

13. On Discrete Residual and Past Measures of Inaccuracy ***Vikas Kumar, *Richa Thapliyal,*** *and* ***H.C. Taneja***

14. Silicon Photonics: A Review ****Ashish Dhiman***

15. Simultaneous Approximation for Sz´asz-Mirakian-Stancu-Durrmeyer Operators Vijay Gupta, ****Naokant Deo*** and Xiaoming Zeng

16. Synthesis of CdS nanocrystals in poly (3-hexylthiophene) polymer matrix: optical and structural studies Vikash Agrawal Kiran Jain, ****Leena Arora*** S. Chand

*** Faculty**
 # Alumni

3-D Localization of an Underwater Sound Source (Pinger) using a Passive SONAR System

Prateek Murgai
Department of Electrical and Electronics Engineering
Delhi Technological University, Shabad Daulatpur, Main Bawana road, New Delhi-110042, India

prateekmurgai64@gmail.com

***Abstract*- The paper introduces a passive sound source localization algorithm using an array of 4 hydrophones (underwater piezoelectric sensors). The algorithm is based on TDOA (Time difference of arrival) estimation employing Cross-correlation and GCC-PHAT(Generalized Cross Correlation using Phase Transform) algorithms .To estimate the position of the pinger intersecting spherical surfaces provide a closed form solution given the time difference of arrival measurements when the distance of the source to any arbitrary reference is unknown.**

I. INTRODUCTION

The problem of passive localization of sound sources using the time difference of arrival from multiple sensors is an active research area in the field of underwater acoustics, radar, bio-acoustics, avalanche detection, unmanned vehicles, etc. Application techniques includes use of hydrophones along a line, surface arrays like a square array and even 3-D arrays like tetrahedron, pyramidal or spherical arrays , Analog to Digital converters (Data Acquisition Cards [5] , TDOA estimations techniques

There are many techniques to estimate the position of a sound source based on energy densities, intensities of signals being received but the most common approach for passive source localization is to use time delays between pair of sensors to define curves of constant time difference or path difference which are hyperbolas of revolution also called as hyperboloids [5], [6] . Intersection of these hyperboloids yields the position of the source. In the next section it will be shown that minimum of 4 sensors are required to locate a source [6]. The solution is typically found numerically since closed form solutions to simultaneous hyperbolic algebraic equations are difficult to find. In most of the applications the source to be located is at a distance much greater than the distance between the sensors, so the direction of the source is given by the asymptotes of the hyperboloids. In 2-D it is given by intersection of 2 lines whereas in 3-D the problem is reduced to the intersection of 3 sets of cones (asymptote lines rotated about axis of each pair of sensors).

One of the basic problems with numerically solving hyperboloid equations is that it is difficult to form a closed form solution, but this is not true for intersecting spheres. This is because hyperboloids are defined as surfaces of constant difference of distances between pair of sensors, while spheres define surfaces of constant distances from single sensor.

Thus keeping in mind about difficulties in solving hyperboloids, we employ the method involving intersection of spheres [1], [2],[3] .

II. HYPERBOLOID EQUATIONS

We start with a set of N hydrophones at positions R_i= (X_i ,Y_i, Z_i). Designate source localization as R_S= (X_S,Y_S,Z_S) we arbitrarily set R_0= (0,0,0) choosing 0^{th} hydrophone as origin.

Each hydrophone receives sound at a time T_i, Each hydrophone pair gives a time difference of arrival as Tij= T_i- T_j and the corresponding range difference $D_{ij} = V^* T_{ij}$ (V is the speed of sound).

Thus

$$\left|\left|R_i - R_s\right| - \left|R_j - RS\right|\right| = D_{ij} \quad (1) .$$

Arbitrarily selecting J=0. We get:-

$$\left|\left|R_i - R_s\right| - \left|R_0 - RS\right|\right| = D_{i0} \quad (2) .$$

Equation (2) represents a hyperboloid with a constant path difference between a specific pair of hydrophones [6] .

Thus it is clear that a pair of hydrophones yields one equation so to localize a source in 3-D we need three equations i.e. a minimum of 4 hydrophones are required to successfully locate the source.

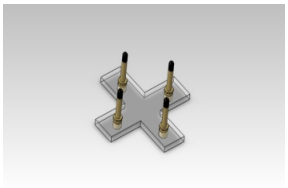


Fig 1. Square array of hydrophone made in Solid Works

Fig 1. Shows a square array of hydrophones which has been used for localization of the pinger [5]. A symmetrical array such as the above helps in noise cancellation while making the time difference of arrival measurements.

III. TIME DIFFERENCE OF ARRIVAL MEASUREMENTS

Two methods of TDOA estimation are presented, they are as follows:

A)Cross Correlation

In signal processing cross correlation is a measure of similarity between 2 signals as a function of time lag between them.

Let there be 2 signals:

S1 (t) = sig1 (t) + n1 (t)

S2 (t) = sig2 (t) + n2 (t)

where sig1 (t) and sig2 (t) are the original signals at each input port and n1 (t) and n2 (t) are the respective noise signals at each port.

For continuous signals S1 (t) and S2 (t) cross correlation is defined as:

$$[S1 * S2](t) = \int_{-\infty}^{+\infty} S1^*(\tau) S2(\tau + t) d\tau$$

where * is the complex conjugate of a function

TDOA is given as follows:

$$TDOA = argmax([S1 * S2](t))$$

As the signals we will be using are sampled discrete signals cross correlation for discrete signals is defined as:

$$(S1 * S2)[n] = \sum_{k=-\infty}^{+\infty} S1^*[k] S2[n + k]$$

where * is the complex conjugate of a function

$$TDOA = argmax((S1 * S2)[n])/Fs$$

where Fs is defined as the sampling rate.

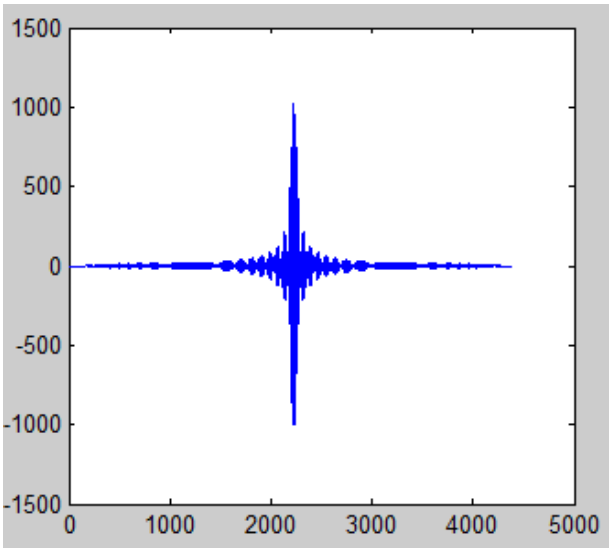


Fig 2. Typical Cross Correlation function simulated in MATLAB

B)GCC-PHAT (Generalized Cross-Correlation using Phase transform)

$$X1(f) = F(S1(t))$$

$$X2(f) = F(S2(t))$$

where F is the Fourier transform of a function.

GCC-PHAT ,[4] is defined as:

$$GPHAT = \frac{X1(f)[X2(f)]^*}{|X1(f)[X2(f)]^*|}$$

where * is the complex conjugate of a function.

$$TDOA = argmax(Rphat(t))$$

and Rphat (t) is defined as:

$$Rphat(t) = F^{-1}(GPHAT)$$

where F^{-1} is the inverse Fourier transform of a function.

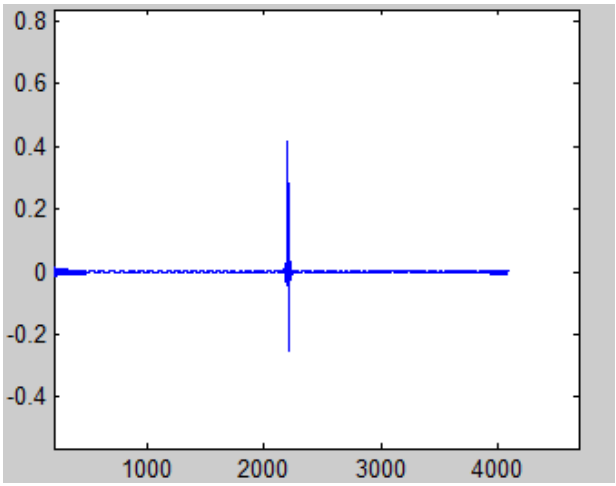


Fig 3. Typical GCC-PHAT simulation in MATLAB

In environments of high levels of reverberation GCC-PHAT helps to improve robustness and accuracy in calculating the time difference of arrival [3], [4]. We can see from the above MATLAB simulation that GCC-PHAT enhances the peak and whitens the region around it, whereas in the Cross Correlation simulation the region around the peak has some disturbance, also the peak is not that well defined as it is in the GCC-PHAT .

Both the methods show similar performance under noise, GCC-PHAT is better in case of added reverberations.

IV. PASSIVE LOCALIZATION USING INTERSECTION OF SPHERES

For four sensors, we can define the path difference from source to sensor *i* and source to sensor *j* as d_{ij} , in terms of absolute distances D_i from source to each of the sensors as follows:

$$d_{ij} = D_i - D_j$$
$$d_{ij} = V * T_{ij} \quad , \qquad (1)$$

V is the speed of sound.

where $T_{ij} = T_i - T_j$,which is also equal to the time difference of arrival.

Take in consideration the distance of the source to the sensors, which give us four basic equations,

$$(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2 = D_i^2 \quad i = 1,4 \qquad (2)$$

where (x_i , y_i , z_i) are the positions of the four sensors and (x_s , y_s , z_s) is the position of the sound source , D_i is the distance of the source to each sensor.
Simplifying equation (2) gives

$$x_i^2 - 2x_s x_i + x_i^2 + y_i^2 - 2y_s y_i + y_s^2 + z_i^2 - 2z_s z_i + z_s^2 = (d_{i4} + D_4)^2 , \quad i = 1,3 \qquad (3)$$

$$x_4^2 - 2x_s x_4 + x_s^2 + y_4^2 - 2y_s y_4 + y_s^2 + z_4^2 - 2z_s z_4 + z_s^2 = D_4^2 \qquad (4)$$

Solving for the distance of source from the origin modifies the above equation as following:

$$R_s^2 = x_s^2 + y_s^2 + z_s^2 \qquad (5)$$

$$R_s^2 = -(x_4^2 + y_i^2 + z_4^2) + D_4^2 + 2x_s x_4 + 2y_s y_4 + 2z_s z_4 \qquad (6)$$

$$R_s^2 = -R_4^2 + D_4^2 + 2x_s x_4 + 2y_s y_4 + 2z_s z_4 \qquad (7)$$

Substituting Rs in (3) into (6)

$$R_i^2 - (d_{i4} + D_4)^2 + D_4^2 - R_4^2 = 2x_i x_s + 2y_i y_s + 2z_i z_s - 2x_4 x_s - 2y_4 y_i - 2z_4 z_i$$

$$i = 1,3 \qquad (7)$$

without loss of generality, we may take sensor 4 at origin ,which gives us (x_4 , y_4 , z_4) = (0 , 0 , 0) .

$$\begin{bmatrix} R_1^2 - d_{14}^2 \\ R_2^2 - d_{24}^2 \\ R_3^2 - d_{34}^2 \end{bmatrix} - 2(x_s^2 + y_s^2 + z_s^2)^{\frac{1}{2}} \begin{bmatrix} d_{14} \\ d_{24} \\ d_{34} \end{bmatrix} = 2 \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \qquad (8)$$

In terms of symbols above equation can be written as:

$$\zeta - 2R_s \alpha = 2\omega X \qquad (9)$$

where

$$\zeta = \begin{bmatrix} R_1^2 - d_{14}^2 \\ R_2^2 - d_{24}^2 \\ R_3^2 - d_{34}^2 \end{bmatrix}$$

$$\alpha = \begin{bmatrix} d_{14} \\ d_{24} \\ d_{34} \end{bmatrix}$$

$$\omega = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$$

$$X = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

From the above equations the source position matrix can be written as following:

$$X = \frac{\omega^{-1}(\zeta - 2R_s \alpha)}{2} \qquad (10)$$

The above equation describing the source stands valid only if ω is a non-singular matrix. The above equation still includes Rs which is an unknown, thus we need to reduce the above equation which has only one set of unknown .To solve this problem we carry out following computation.

$$R_s = (X^T X)^{1/2} \qquad (11)$$

Substituting (10) into (11) yields the following equation:

$$R_s^2 (4 - 4\alpha^T (\omega^{-1})^T \omega^{-1} \alpha) + R_s (2\alpha^T (\omega^{-1})^T \omega^{-1} \zeta + 2\zeta^T (\omega^{-1})^T \omega^{-1} \alpha) - (\zeta^T (\omega^{-1})^T \omega^{-1} \zeta) = 0$$

In the above equation the resultant of terms in the brackets are 1x1 matrices, i.e. they represent the scalar values of the 1x1 matrices. Solving the quadratic equation gives us the following results:

$$R_s = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where

$$a = \left(4 - 4\alpha^T(\omega^{-1})^T\,\omega^{-1}\alpha\right)$$

$$b = \left(2\alpha^T(\omega^{-1})^T\,\omega^{-1}\zeta + 2\zeta^T(\omega^{-1})^T\omega^{-1}\alpha\right)$$

$$c = -(\zeta^T(\omega^{-1})^T\omega^{-1}\zeta)$$

We consider only the values for which R_s (range) is greater than zero .If we get two positive values of R_s , the actual value of the range is selected based upon some criterion separating the two values ,usually the difference between the two values is large so that they can be separated easily. Once the range is achieved we can substitute this value in (10) to achieve the source position or coordinates w.r.t a reference origin i.e. the sensor 4.

V. MATLAB SIMULATION RESULTS

In this section we construct a simulation model in MATLAB to forecast the system performance. We assume a 3-D microphone array consisting of 4sensors as displayed in the figure below.

To simulate a localization situation we do perform the following task

Let the pinger position be (20, 16, 5).
Let the sensor to sensor spacing = 28 cm = 0.28m =d.

Sensor 1 position = (d, 0, 0)
Sensor 2 position = (0, d, 0)
Sensor 3 position = (0, 0, d)
Sensor 4 position = (0, 0, 0)

Speed of Sound = 1484 m/s
Path difference between sensor 4 to source and sensor 1 to source = 0.2291083 m

Path difference between sensor 4 to source and sensor 2 to source = 0.15208991m

Path difference between sensor 4 to source and sensor 3 to source = 0.046521 m

TDOA1 = (0.2291083/1484) = 1.543e-4
TDOA2 = (0.15208991/1484) = 1.024e-4
TDOA3 = (0.046521/1484) = 0.3134e-4

N1=Number of delayed samples 1 = TDOA1/Fs ~ 46 samples
N2=Number of delayed samples 2= TDOA2/Fs ~ 31 samples
N3= Number of delayed samples 3=TDOA/Fs ~ 9 samples

where Fs is the sampling rate .

To simulate signals we have taken chirps as our input signals. These chirps are generated within MATLAB itself .These chirps have an initial frequency f0=22000 Hz and final frequency f1= 27000Hz with a central frequency of 25000Hz.

We have assumed the speed of sound underwater is approximately 1484m/s and sampling rate is 200000KS/s

The chirps which are assigned to sensor 1, 2, and 3 are delayed with N1, N2, N3 samples respectively w.r.t the chirp assigned to the sensor 4.

As we have our 4 input signals in the form of chirps we can reverse engineer by first applying TDOA methods and then apply the Spherical Intersection algorithm to estimate position. The results of the simulation for both the TDOA methods are stated below.

Pinger Position (meters)			Cross-Correlation Estimated Position			|%Error|		
xs	ys	zs	xc	yc	zc	x	y	z
24	16	5	23.75	16.19	4.77	1	1.18	4.6
40.2	10.5	6.3	40.23	10.5575	6.095	0.07	0.54	3.2
16.6	5.7	3.9	16.47	5.80	3.97	0.78	1.75	1.18
10.6	3.5	3.1	10.69	3.413	3.209	0.85	2.48	3.51
33.9	10.3	6.4	33.76	10.3160	6.502	0.41	0.16	1.6

TABLE 1: Cross-Correlation Results

Pinger Position (meters)			GCC-PHAT Estimated Position			|%Error|		
xs	ys	zs	xc	yc	zc	x	y	z
24	16	5	23.75	16.1	4.75	1	0.6	5
40.2	10.5	6.3	40.22	10.55	6.09	0.05	0.47	3
16.6	5.7	3.9	16.47	5.82	3.93	0.78	2.1	0.77
10.6	3.5	3.1	10.65	3.39	3.19	0.47	3.14	2.9
33.9	10.3	6.4	33.74	10.31	6.49	0.47	0.01	1.4

TABLE II: GCC-PHAT Results .

V. HARDWARE IMPLEMENTATION OF PASSIVE SONAR SYSTEM.

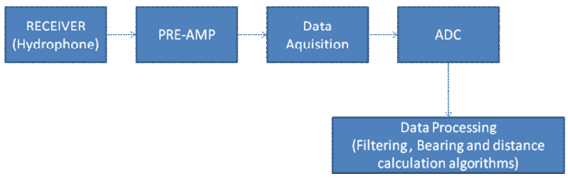


Fig 4. Underwater localization task flowchart.

The A/D being used is a NI-DAQ PCI 4462 which is a 24-Bit, 204.8KS/s data acquisition module which has 4 input ports. The hydrophones being used are TC-4013 which have a frequency range of 1Hz to 170 KHz and provide omnidirectional sensitivities. For onboard processing a Single Board Computer (SBC) or a laptop is used.

ACKNOWLEDGEMENT

The author gratefully acknowledge Ms. Hemlata Choudhary, Mr. .Vidush Saxena for their contribution to this journal paper and for making the concepts more realizable. The author would also like to appreciate the support of DTU-AUV(Delhi Technological University Autonomous Underwater Vehicle) team for providing hardware inputs of National Instruments DAQ (Data Acquisition Card) , Hydrophones and a SBC .

VI. CONCLUSION

In this paper, we proposed a method for localization of an acoustic source using an array of hydrophones (passive system).Technology based on TDOA method is studied and the theory and simulation of 2 such methods namely cross-correlation and GCC-PHAT are presented. Further an algorithm based on spherical intersection is used to estimate to position of the pinger along with range from the sensor array. Simulated results show that the maximum percentage error in (x, y, z) coordinates is respectively and the maximum percentage error in x_s is 1%, in y_s is 3.14% and in z_s is 5%. We also observe that the error in position due to both TDOA methods is quite similar as we have not created a situation of artificial reverberation, thus both methods work equally good. In conclusion, the implemented method in the paper can be applied to source localization and achieve fairly good location accuracy.

REFERENCES

[1] H.C. SCHAU and A.Z. ROBINSON, "Passive Source Localization employing intersecting spherical surfaces from Time of Arrival Differences " , IEEE Transaction , Acoustics ,Speech , Signal Processing ,Vol ASSP-35 , No .8 ,AUGUST 1987 .

[2] Brian .G .Ferguson, Member, IEEE, Kam W. Lo, Senior Member IEEE ,and Ranjit A .Thuraisingham , " Sensor Position Estimation and Source Ranging in a Shallow Water Environment " , IEEE Journal of Oceanic Engineering ,Vol 30, No. 2 , April 2005 .

[3] Hamid Reza Abutalebi and Hossein Momenzadeh ," Performance Improvement of TDOA Based Speaker Localization in Joint Noisy Reverberant Conditions , EURASIP Journal on Advances in Signal Processing Volume 2011 ,Article ID 621390 .

[4] C Knapp, G Carter, "The generalized correlation method for estimation of time delay," IEEE Transaction of Acoustics Speech and Signal processing, Vol 24(4), pp. 320-327, 1976.

[5] Sinha, R.K. Delhi Technological. University. (Formerly Delhi College. of Eng.), Delhi, India , Ahmed, S. ; Tutwani, M. ; Bahuguna, P. ; Kumra, R. ; Agarwal, V. ; Saxena, V. ; Mittal, V. ; Gupta, P. , " Vehicle for Automation Research and Underwater Navigation " .

[6] Leo Singer, University of Maryland, "SONAR cookbook 'An Underwater Primer for TORTUGA –II , " July 28 ,2008 .

An Optimal Edge Detection Using Gravitational Search Algorithm

Om Prakash Verma, Rishabh Sharma, Manoj Kumar, and Neetu Agrawal

***Abstract*—Edge detection is a fundamental tool used in most image processing applications to obtain information from the image as a precursor step to feature extraction and object segmentation. Gravitational search algorithm (GSA) is a new population-based search algorithm inspired by Newtonian gravity. Algorithm uses the theory of Newtonian gravity and its searcher agents are the collection of masses. Masses attract each other by way of gravity force, and this force causes a global movement of all objects towards the objects with heavier masses. In the proposed approach the edges are detected by the local variation in intensity values and the movement of agents is computed using gravitational search algorithm. The proposed approach is able to detect the edge pixel in an image up to a certain extent. The technique can be further extended for finding more edge pixels by modifying the gravitational search algorithm.**

***Index Terms*—Gravitational search algorithm, edge, intensity, Fitness function, optimization.**

I. Introduction

Edge detection is a fundamental of low-level image processing and good edges are necessary for higher level of image processing [1]. Edges are one of the most important visual clues for interpreting images [2]. The process of edge detection reduces an image to its edge details that appear as the outlines of image objects that are often used in subsequent image analysis operations for feature detection and object recognition.

A very important role is played in image analysis by what are termed feature points, pixels that are identified as having a special property. Feature points include edge pixels as determined by the well-known classic edge detectors of Sobel [3], Prewitt [4], Kirsch [5], Canny [6] etc. Russo [7], [8] and Russo and Ramponi [9], designed fuzzy rules for edge detection. Such rules can smooth while sharpening edges, but require a rather large rule set compared to simpler fuzzy methods. Abdallah and Ayman [10] introduced a fuzzy logic reasoning strategy for the edge detection in the digital images without determining a threshold.

Manuscript received December 12, 2012; revised March 27, 2013.

Om Prakash Verma is with the Department of Information Technology, Delhi Technological University (formally DCE), New Delhi, DL 110042, India (e-mail:opverma.dce@gmail.com).

Rishabh Sharma is with the Ericsson India Global Services Pvt. Ltd, Noida, India (e-mail:rishabhs.32@gmail.com).

Manoj Kumar is with the Department of Computer Engineering, Delhi Technological University (formally DCE), New Delhi, DL 110042, India (e-mail:mkg1109@rediffmail.com).

Neetu Agrawal is with the Department of Electronics and Communication, Ambedker Institute of Technology, New Delhi, DL 110031, India (e-mail:neetu24@gmail.com).

Over the last decades, there has been a growing interest in algorithms inspired by the behaviors of natural phenomena. It is shown by many researchers that these algorithms are well suited to solve complex computational problems. Genyun Suna *et al.* [11] have introduced an edge detection algorithm based on the law of universal gravity in 2007. This algorithm assumes that each image pixel is a celestial body with a mass represented by its grayscale intensity. Verma *et al.* [12] have also developed a novel fuzzy system for edge detection in noisy image using bacterial foraging. Another evolutionary technique known as Particle Swarm Optimization (PSO) [13] employs a swarming in which the movements of the particles are guided by the swarm's local best position as well as global best position in the required search-space. Verma *et al.* [14] have also developed a new approach for edge detection using fuzzy derivative and Ant Colony Optimization. (ACO) algorithm to reduce the discontinuities presented in the image filtered by Sobel operator. Recently Verma *et al.* [15] proposed a new optimal approach for edge detection using universal law of gravity and ant colony optimization. In this approach, the theory of universal gravity is used to calculate the heuristic function which guides the ant towards edge pixels.

Edge detection aims to localize the boundaries of objects in an image and is a basis for many image analysis and machine vision applications. Conventional approaches to edge detection are computationally expensive because each set of operations is conducted for each pixel. In conventional approaches, the computation time quickly increases with the size of the image.

Gravitational search algorithm [16] is an optimization algorithm inspired by Newtonian gravity. Masses cooperate using a direct form of communication, through gravitational force of attraction. Each mass presents a solution, and the algorithm is navigated by properly adjusting the gravitational and inertia masses. The lighter masses tend to get attracted towards heaviest mass. The heavier mass presents an optimum solution in the search space.

In this paper, GSA is used to tackle the edge detection problem in optimal manner. Local variation of image intensity value is used to detect edge pixels while the movement of agents are computed using the gravitational search algorithm.

The rest of the paper is organized as the follows. The gravitational search algorithm is briefly reviewed in Section II. The algorithm of the proposed edge detector is presented in Section III. The experimental results are given in Section IV and conclusions are drawn in Section V.

DOI: 10.7763/LNSE.2013.V1.33

II. Gravitational Search Algorithm

In GSA, Newtonian laws of gravity are applied to find the optimum solution by a set of agents called masses [16]. In GSA, each mass (agent) has four characteristics namely; 1) position, 2) inertial mass, 3) active gravitational mass, and 4) passive gravitational mass. The position of the mass corresponds to a solution of the problem, and the fitness function is used to determine the gravitational and inertial masses.

The GSA algorithm is mainly comprises of the following steps:

1) Identification of search space.
2) Initialization.
3) Agent evaluation using fitness function.
4) Update $G(t)$, $best(t)$, $worst(t)$ and $M_i(t)$ for i=1,2,..*N*.
5) Calculation of the total force in all possible directions.
6) Acceleration and velocity calculations.
7) Updating agents' position.
8) Repeat steps 3 to 7 until the stop criterion is reached.
9) End

Consider a system with *N* masses in which position of the i^{th} mass is defined as follows [16]:

$$X_i=(x_i^1,\ldots,x_i^d,\ldots,x_i^n) \quad for\, i=1,\ 2,\ \ldots,\ N \tag{1}$$

where x_i^d is position of the i^{th} mass in the d^{th} dimension and *n* is dimension of the search space.

Mass of each agent is calculated after computing current population's fitness as follows:

$$m_i(t)=\frac{fit_i(t)-worst(t)}{best(t)-worst(t)}, \tag{2}$$

$$M_i(t)==\frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \tag{3}$$

where $M_i(t)$ and $fit_i(t)$ represent the mass and the fitness value of the agent *i* at time *t*, respectively. In case of minimization problem, $worst(t)$ and $best(t)$ are defined as follows:

$$worst(t)=\max_{i\varepsilon\{1,\ldots,N\}} fit_i(t) \tag{4}$$

$$best(t)=\min_{i\varepsilon\{1,\ldots,N\}} fit_i(t) \tag{5}$$

The net force, applied by mass *j* on mass *i*, $f_{ij}(t)$, is calculated as:

$$f_{ij}(t)=G(t)\frac{M_i(t)M_j(t)}{R_{ij}(t)+\varepsilon}(x_j-x_i) \tag{6}$$

here x_i is the position vector of the i^{th} agent, ε is a small threshold, and $G(t)$ is the gravitational constant, initialized at the beginning of the algorithm. The $G(t)$ is decreased with time to control the search accuracy. $R_{ij}(t)$ is the Euclidian distance between two masses *i* and *j*. Using the Newton's second law of motion, the total gravitational acceleration on the i^{th} agent, is calculated as:

$$a_i(t)=G(t)\sum_{j=1}^{n} rand_j \frac{M_j(t)}{R_{ij}(t)+\varepsilon}(x_j-x_i) \tag{7}$$

where r and $_j$ is a random number within the interval [0, 1], considered to add some stochastic behavior to the acceleration. The velocity and position of the agents are updated as:

$$v_i(t+1)=rand_i\ v_i(t)+a_i(t) \tag{8}$$

$$x_i(t+1)=x_i(t)+v_i(t+1) \tag{9}$$

k_{best} is the set of first K agents with the best fitness value and biggest mass, which is a function of time , initialized to k_0 at the beginning and decreasing with time. Here k_0 is set to N (total number of agents) and is decreased linearly to 1.

III. The Proposed Image Edge Detection Approach

The proposed approach using the gravitational search algorithm leads to a minimal set of input data to be processed. The proposed GSA-based image edge detection approach aims to utilize a set of agents called masses to move on a 2-D image under the influence of gravity.

Algorithm

STEP 1: Initialize the positions of *N* agents by randomly distributing them on an image *I*. Initialize the value of the gravitational constant $G(t_0)$

STEP 2: Set the value of k_{best} equal to the number of agents

STEP 3: While ($k_{best} \neq 1$)

STEP 4: For each agent:

1) *Calculate the total force applied on a single agent by all the other agents in the search space:* In this approach total force on an agent is equal to the sum of the force exerted by all the agents individually on a single agent. Force exerted by an agent *i* at location (*x*, *y*) on agent *j* at location (x_1 , y_1) is calculated using the Eq. 6, Where $M_i(t)$ and $M_j(t)$ are the grey value of the pixels located at (*x*, *y*) and (x_1 , y_1). $R_{ij}(t)$ is the Euclidian distance between two agents *i* and *j* which is calculated using the eq.

$$\| \vec{R}_{ij} \| = \sqrt{(x_1-x)^2+(y_1-y)^2} \tag{10}$$

2) The total force exerted on an agent is calculated as:

$$f_i(t)=\sum_{j=1,\, j\neq i}^{N} rand_j\, f_{ij}(t) \tag{11}$$

3) *Calculate the acceleration and velocity of an agent:* In this approach the acceleration of each individual agent is calculated using the Eq. 7. Using this acceleration, velocity of the individual agent is calculated using the Eq.8.
4) *Update the agent's position:* Each individual agent's position is updated using Eq. 9. where velocity is calculated in the previous step.
5) *Updating* $M_i(t)$, $G(t)$ *:* As each individual agent's position is updated. The grey value of the pixel on which it lands on becomes its new mass. The heavy masses which correspond to good solutions move more slowly than lighter ones. The $G(t)$ is updated as:

$$G(t)=G(t_0)\left(\frac{t_0}{t}\right)^{\beta} \quad \beta<1 \tag{12}$$

6) *Fitness evaluation of agents:* In the proposed approach edges are detected using the local variations of image intensity value. For each agent we consider a 3 × 3 neighborhood (Fig. 1). Information required to find out whether a pixel is an edge pixel or not is determined by the local statistics at that position [17]:

$$f_{fitness}=\frac{V_c(I_{x,y})}{V_{\max}} \tag{13}$$

where $I_{x,y}$ is the intensity value of the pixel at (*x,y).* $V_c(I_{x,y})$ is a function that operates on the local group of pixels (Fig. 1) around the pixel *(x, y)* . It depends on the variation of the intensity values on the local group, and is given by [17]

$$V_c(I_{x,y})=|I_{x-1,y-1}-I_{x+1,y+1}|+|I_{x-1,y}-I_{x+1,y}|+ |I_{x-1,y+1}-I_{x+1,y-1}|+|I_{x,y-1}-I_{x,y+1}| \tag{14}$$

$V_{\max}$ is the maximum intensity variation in the whole image and serves as a normalization factor.

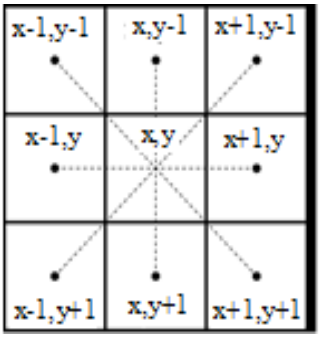


Fig. 1. A local configuration for computing the intensity variation at (*x, y*).

STEP 5: End for loop.

STEP 6: Update k_{best} .

Decreasing the value of k_{best} (no. of agents) by removing agents that did not lead to edge pixels .

The flowchart of the proposed algorithm is shown in Fig. 2.

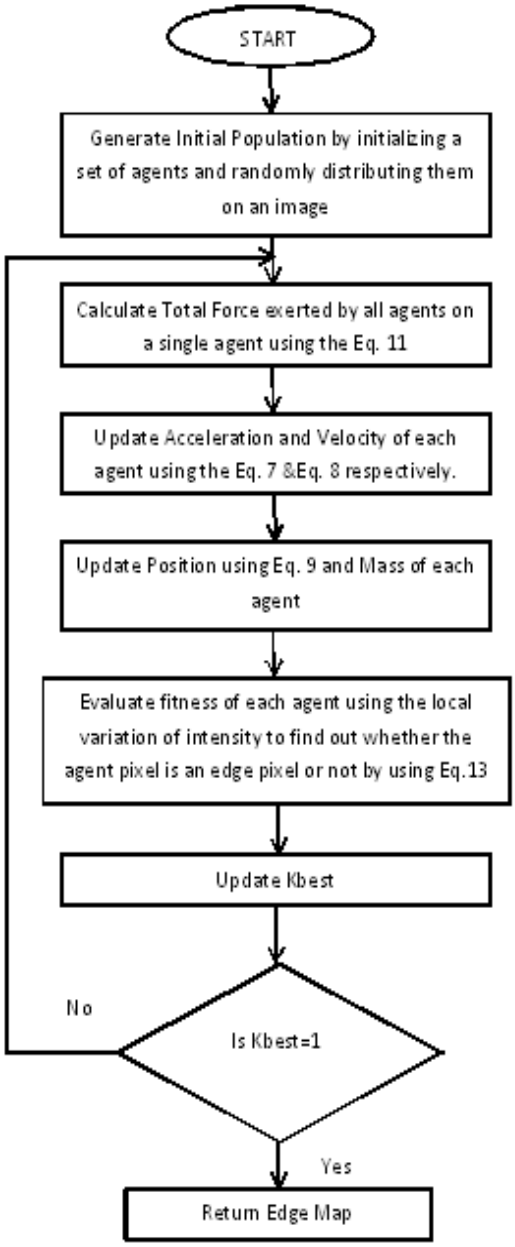


Fig. 2. Proposed algorithm.

IV. Experimental Results

Conventional approaches to edge detection are computationally expensive because each set of operations is conducted for each pixel. The proposed approach is more focused on the optimization of the edge detection problem. Furthermore, various parameters of the proposed approach are set as follows [16]:

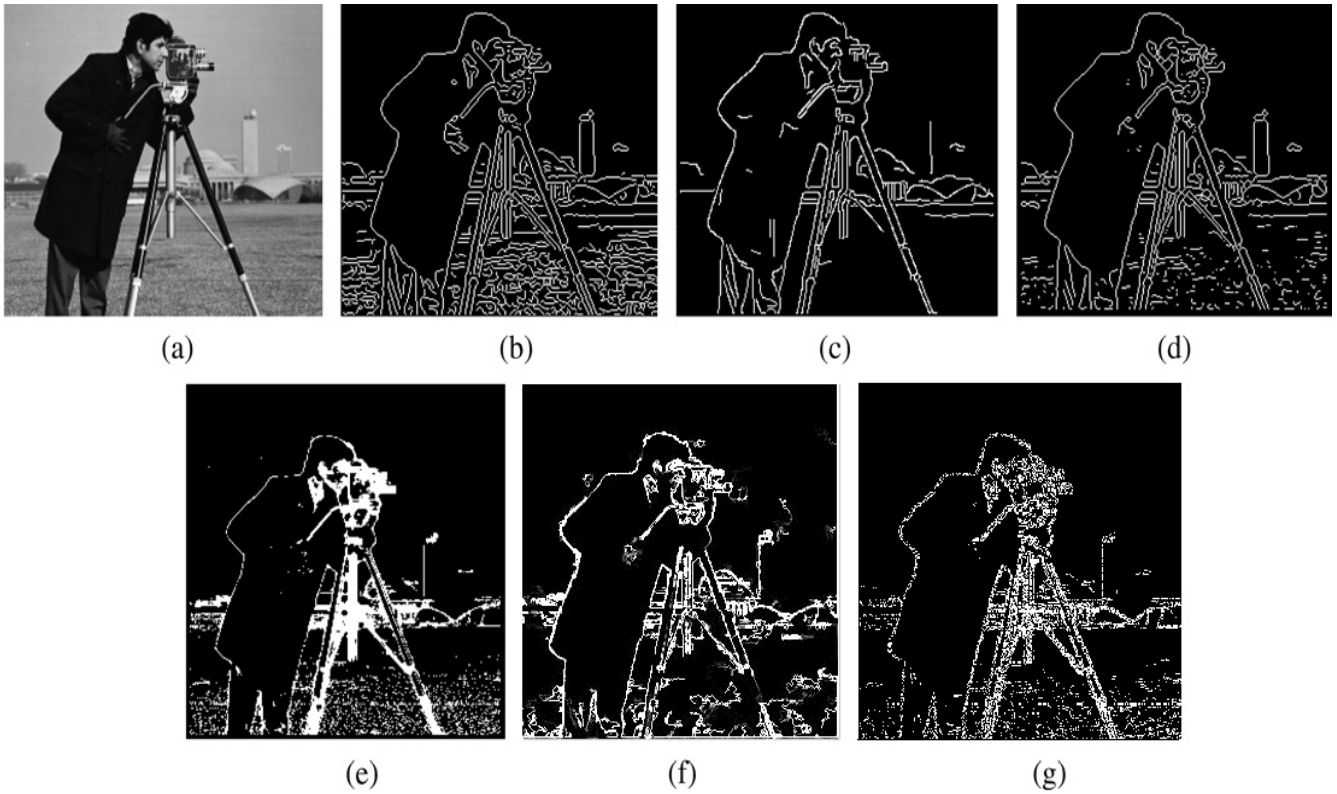


Fig. 3. (a) Original Cameraman image (b) Canny Edge Detector (c) Edison Edge Detector (d) Rothwell Edge Detector (e) SUSAN Edge Detector (f) Edge detection based on the theory of universal gravity and Ant colony optimization (threshold: default) and (g) The proposed approach.

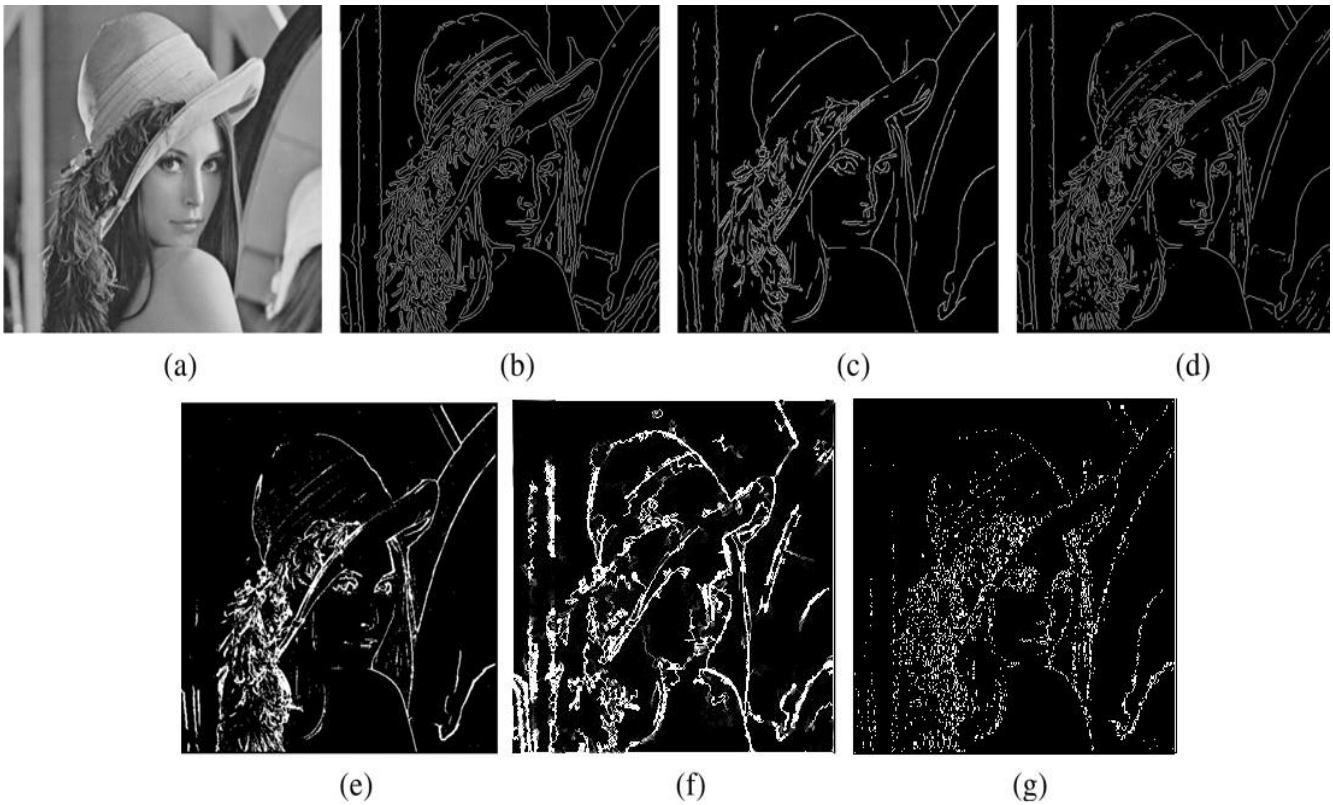


Fig. 4. (a) Original lena image (b) Canny Edge Detector (c) Edison Edge Detector (d) Rothwell Edge Detector (e) SUSAN Edge Detector (f)Edge detection based on the theory of universal gravity and ant colony optimization (threshold: default) and (g) The proposed approach.

$s = \sqrt{M\ 1 \times M\ 2}$ (the total number of agents), Ω =8-connectivity neighborhood, $G(t_0)$ =1, β =0.2, *N=50*, $rand_j$ =rand function (leads to a value between 0 and1).

The performance of the proposed technique is compared against that of the traditional edge detectors such as Canny, Edison, Rothwell, and SUSAN. The performance is also compared with the edge detector using universal law of gravity and ant colony optimization [15]. In terms of number of edges, the traditional operators performs slightly better than the proposed approach but the proposed approach gives better results than the edge detector using the universal law of gravity and ant colony optimization. The proposed approach leads to less time consumption and less memory exaustive

then the traditional operators and thus aptly focusing on the optimization concerns of the edge detection problem.

For images in Fig. 3 and Fig. 4, the captions are as follows: (a) the original image, (b) Canny Edge Detector, (c) Edison Edge Detector, (d) Rothwell Edge Detector, (e) SUSAN Edge Detector, (f) edge detection based on the theory of universal gravity and ant colony optimization (threshold: default) and (g) the result of the proposed approach. The value of threshold for the standard Canny, Edison, Rothwell, and SUSAN operator is selected as default value which gives the ideal edge map. The traditional edge detectors are implemented using the MATLAB toolbox. The proposed method and edge detection using universal law of gravity and ant colony optimization is also coded in the MATLAB.

It is evident from the results that our method finds meaningful edges in most images but is partially successful in presenting the complete edge map.

V. Conclusion

In this paper, a new GSA based image edge detection approach has been successfully developed. The proposed approach uses local variations of image intensity in-order to detect edges. The movement of the agents is dependent on the gravitational search algorithm. The proposed approach leads to a minimal set of input data to be processed thus making the process much faster and memory-efficient then the edge detection algorithm [15].

References

[1] M. B. Ahmad and T. S. Choi, "Local Threshold and Boolean Function Based Edge Detection," *IEEE Transactions on Consumer Electronics*, vol. 45, no 3, pp. 674 -679, August 1999.

[2] E. Gose, R. Johnsonbaug, and S. Jost, *Pattern Recognition and Image Analysis*, Englewood Cliffs, NJ: Prentice-Hall, 1996, pp. 298.

[3] I. Sobel, "Camera Models and Perception," Ph.D. thesis, Stanford University, Stanford, CA, 1970.

[4] J. M. S. Prewitt, "Object enhancement and extraction," in B. S. Lipkin and A. Rosenfeld (Eds.), *Picture Analysis and Psychopictorics*, New York, NY: Academic Press, pp. 75-149, 1970.

[5] R. Kirsch, "Computer determination of the constituent structure of biological images," *Computers and Biomedical Research*, vol. 4, pp. 315-328, 1971.

[6] J. Canny, "Computational Approach to Edge Detection," *IEEE Transaction on Pattern Analysis and Machine intelligence*, no. 6, pp. 679-698, 1986.

[7] F. Russo, "A new class of fuzzy operators for image processing," in *Proc. IEEE Int. Conf. Neural Netw.*, vol. 1, pp. 815–820, 1993.

[8] F. Russo, "A user-friendly research tool for image processing with fuzzy rules," in *Proc. the First IEEE International Conference on Fuzzy Systems*, San Diego, 1992, pp. 561–568.

[9] F. Russo and G. Ramponi, "Fuzzy operator for sharpening of noisy images," *IEE Electron. Lett.*, vol. 28, pp. 1715–1717, 1992.

[10] A. A. Abdallah and A. A. Ayman, "Edge detection in digital images using fuuzy logic techniques," *World Academy of Sci. Eng. Technol.*, vol. 51, pp. 178–186, 2009.

[11] G.-Y. Suna, Q.-H. Liua, Q. Liua, C.-Y. Jib, and X.-W. Lia, "A novel approach for edge detection based on the theory of universal gravity," *Pattern Recognition*, vol. 40, no. 10, pp. 2766-2775, 2007.

[12] O. P. Verma, M. Hanmandlu, and S. A. K. Dhruv, "A Novel Fuzzy Ant System for Edge Detection," *IEEE/ACIS 9th International Conference on Computer and Information Science*, pp. 228 – 233, 2010.

[13] M. Setayesh, M. Zang, and M. Jonhnton, "A new homogeneity-based approach to edge detection using PSO," in *Proc. 24th International Conference Image and Vision Computing*, New Zealand (IVCNZ), 2009, pp. 231-236.

[14] O. P. Verma, M. Hanmandlu, P. Kumar, and S. Srivashtav, "A novel approach for edge detection using ant colony optimization and fuzzy derivative technique," *Advance Computing Conference, IEEE international*, pp. 1206–1212, 2009.

[15] O. P. Verma and R. Sharma, "An optimal edge detection using universal law of gravity and Ant colony algorithm," *World congress on communication and information technologies (WICT)*, 2011.

[16] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "GSA: A Gravitational Search Algorithm," *Information Sciences*, vol. 179, pp. 2232–224, 2007.

[17] H. Nezamabadi-pour, S. Saryazdi, and E.Rashedi, "Edge Detection Using Ant Algorithms," *Soft Computing*, vol. 10, pp. 623-628, 2006.

Om Prakash Verma is the Head of the Department of Information Technology, Delhi Technological University (formally DCE) having experience of more than 20 years in the field of teaching. Mr. Om Prakash Verma specializes in Image Processing, Signal Processing, Soft Computing and Artificial Intelligence. He has graduated from MNIT Jaipur and did his masters from Indian Institute of Technology Delhi (IITD). He has done his Phd from Delhi University. His work has been published in more than 40 journals and conferences proceedings. Under his guidance, more than 25 students have completed their Masters.

Kusum Arora, Dr. S.K. Agarwal, Dr. Narendra Kumar, Dharam Vir / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 3, Issue 3, May-Jun 2013, pp.821-826

Analysis of Power Flow Control in Power System Model using Thyristor Controlled Series Capacitor (TCSC)

Kusum Arora*, Dr. S.K. Agarwal **, Dr. Narendra Kumar *, Dharam Vir******

*(Department of Electronics Engineering, YMCA University of Science & Technology, Faridabad, India)
** (Department of Electronics Engineering, YMCA University of Science & Technology, Faridabad, India)
*** (Department of Electrical Engineering, Delhi College of Engineering, and New Delhi, India)
**** (Department of Electronics Engg, YMCA University of Science & Technology, Faridabad, India)

ABSTRACT

The characteristic of a thyristor controlled series compensator (TCSC) is usually defined by the overall reactance of the device versus the firing angle of the TCR that is connected in parallel with a fixed capacitor. We represents a model control circuit by using microcontroller to generate trigger pulses to fire the gate of the thyristors according to the set firing angle for required output. It brings out the operation of TCSC along with numerical equations. It also gives impedance characteristics curve of a TCSC device and specifies the range of inductance and capacitance region and also finds a suitable value of inductance and capacitance. We analyze the different waveforms in the capacitive as well as inductive region of TCSC, optimal setting of TCSC is determined through the software code written in MATLAB.

Keywords - *Firing angle, Thyristor Controlled Series Compensator (TCSC), Optimal Power Flow, Power system analysis, Zero Crossing Detector*

I. INTRODUCTION

Modern electric power utilities are facing many challenges due to ever-increasing complexity in their operation and structure. In the recent past, one of the problems that got wide attention is the power system instabilities. With the lack of new generation and transmission facilities and over exploitation of the existing facilities make these types of problems more imminent in modern power systems. With the increased loading of transmission lines, the problem of transient stability after a major fault can become a transmission power limiting factor. The use of fixed capacitors to improve transient stability is well known. However, fixed capacitors in series with the transmission lines may cause sub synchronous resonance that can lead to electrical instability at oscillation frequencies lower than the normal system frequency. So there exists the need for switched series capacitors [1]. The switching sequences are determined, based on the sign reversal of a 'decision function', which is derived from the equal-area criterion and requires knowledge of the final steady state. This method is very complicated and later in the author himself adopted a far simpler and more practical control approach .The control law adopted was to insert and bypass a series capacitor based on the speed deviation. At the time the above ideas in and were proposed, dynamically varying series compensation could only be achieved using capacitor banks switched in and out with mechanical circuit breakers; however such devices were in practice too slow and unreliable for high speed use. So these ideas remained on hold for nearly two decades until the evolution of FACTS devices. Flexible AC transmission System (FACTS) controllers use thyristor switching devices to provide greater control, speed and flexibility of ac transmission systems [2]. The Thyristor Controlled Series Compensator (TCSC) is a second generation FACTS controller capable of providing fast variable compensation. This paper focuses on the variable impedance capability of TCSC for enhancing the transient stability. It Mitigate sub synchronous resonance ,Increase power transfer, Optimize power flow between the lines, Reduce system losses, Increase transient stability, Increase dynamic stability, Control current, Reduce damping oscillations, Increase voltage stability, Fault current limiting [1] [3]. There exists a class of control schemes for transient stability enhancement using TCSC. In this paper, transient stability control model scheme, namely: microcontroller based control is done for study. The controller is a simple speed deviation based controller and can provide a drastic improvement in transient stability. In addition to the transient stability enhancement, this controller provides power oscillation damping also. We calculate the various resonance points in the impedance characteristics of the TCSC with respect to various firing angle. We analyses and design the TCSC model using ORCAD and MATLAB. The rest of the paper as follows. Section II Literature Survey, Section III Benefits of TCSC, Section IV modeling of the TCSC, Section

V Analysis of steady state transient behavior the implemented TCSC circuit and finally conclusion and future work in section VI.

II. LITERATURE SURVEY

The Equation of reactive power implies that impedance of series compensating capacitor cancels the portion of the actual line reactance & thereby effective transmission impedance is reduced.

$$Q = I^2 X_C = \frac{2V^2 k(1-COS\delta)}{(1-k)^2}$$

As if the line was physically shortened or we can say that in order to increase the current in the given circuit series impedance of the actual physical line, the voltage across this impedance must be increased. This can be accomplished by an appropriate series connected circuit element such as capacitor, the impedance of which produces a voltage opposite to the prevailing voltage across the series line reactance and thus causing later voltage to increase .So Series compensator provides desired series compensation voltage when and where required if installed [4] [5].

According to S.Meikandasivam, Rajesh Kumar Nima and Shailendra Kumar Jain in their paper "Beaviour Study 0f TCSC presented in World Academy of Science Engineering and technology 45 2008" World's first 3 phase [5], 2 X 165 MVAR, TCSC was installed in 1992 in Kayenta substation, Arizona. It raised the transmission capacity of transmission line by 30%, but it was soon realized that the device is also a very effective means for providing damping of electromechanical power oscillations. A third possible application of TCSC emerged from the onsite observations that it can provide series compensation without causing the same risk for sub-synchronous resonance (SSR) a fixed series capacitor. World's first TCSC for subsynchronous resonance (SSR) mitigation was installed in Stode, Sweden in 1998, by ABB. Specifically this period makes a valiant period for TCSC and makes the researchers to turn on to TCSC And while selecting inductance, X_L should be sufficiently smaller than that of the capacitor X_C to get both effective inductive and capacitive reactance across the device. Also X_L should not be equal to X_C value; or else a resonance develops that result in infinite impedance an unacceptable condition. Note that while varying X_L (α), a condition should not allow to occur X_L (α) = X_C.

Hailian Xie and Lennart Angquist designed a new control scheme instead of traditional firing angle control scheme called as the Static Voltage Reversal Control scheme. Following that scheme the simulation results are analyzed using real time simulator.

T.Venegas and C.R. Fuerete Esquivel report on large scale power flow studies. The ASC model are developed in phase coordinates and incorporated into an existing Newton Raphson power flow algorithm and analysis is done on both balanced and unbalanced power network operating conditions.

Dragan Jovcic, member IEEE and G.N.Pillai presents an analytical, linear, state-space model of TCSC. First a simplified fundamental frequency model of TCSC is proposed and the model results are verified. Using the nonlinear TCSC segment, a simplified nonlinear state space model is derived where frequency of dominant TCSC complex poles shows linear dependence on the firing angle. The nonlinear element is linearized and linked with the AC network model along with TCSC model and is implemented on MATLAB and verified on EMTDC/ PSCAD in frequency and Time domain for a range of operating conditions.

III. OPERATION AND ADVANTAGES OF VARIOUS MODES OF OPERATION, AND BENEFITS OF TCSC

A. CONCEPT OF TCSC:

Concept behind TCSC is to decrease or increase overall effective series transmission impedance from sending end to the receiving end so as to control the transmission of power and the current in the reactor can be controlled from zero to maximum by the method of firing delay angle. Closure of the thyristor valve is delayed w.r.t. peak of the applied voltage in each half cycle thus duration of the current conduction interval is controlled [6].

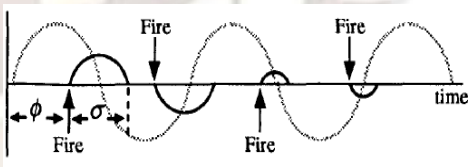


Figure 1. Various ware form shows the operation of thyristor with various firing and conduction angles.

There are three modes of operation of TCSC depending upon the firing angle of the pulses fed to the thyristor.

- Thyristor blocked mode
- Thyristor bypassed mode
- Vernier operating mode

B. Thyristor blocked Operating Mode:

When the thyristor valve is not triggered and the

TCSC is operating in blocking mode. In this mode, the TCSC performs like a fixed series capacitor

C. Thyristor bypass Operating mode:

In bypass mode the thyristor valve is triggered continuously and the valve stays conducting all the time; so the TCSC behaves like a parallel connection of the series capacitor with the inductor, Ls in the thyristor valve branch. In this mode, the resulting voltage in the steady state across the TCSC is inductive and the valve current is somewhat bigger than the line current due to the current generation in the capacitor bank. For practical TCSC's with ratio (X_L/X_C) between 0.1 to 0.3 ranges, the capacitor voltage at a given line current is much lower in bypass than in blocking mode. Therefore, the bypass mode is utilized as a means to reduce the capacitor stress during faults [7] [9].

D. Vernier Operating Mode:

In Vernier control the TCSC dynamics are varied continuously by controlling the firing angle. The firing angle is possible from 0^o to 90^o for each half cycle when it is generated from the zero crossing of the line current hence divided into two parts:

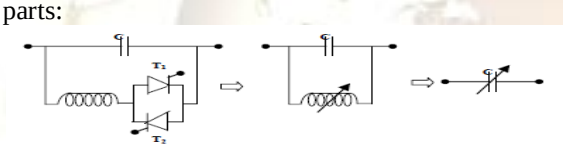


Figure 2. Equivalent circuit of TCSC in Vernier Mode

1) Capacitive Boost mode:

In capacitive boost mode a trigger pulse is supplied to the thyristor having forward voltage just before the capacitor voltage crosses the zero line, so a capacitor discharge current pulse will circulate through the parallel inductive branch. The discharge current pulse adds to the line current through the capacitor and causes a capacitor voltage that adds to the voltage caused by the line current. The capacitor peak voltage thus will be increased in proportion to the charge that passes through the thyristor branch. The fundamental voltage also increases almost proportionally to the charge. From the system point of view, this mode inserts capacitors to the line up to nearly three times the fixed capacitor. This is the normal operating mode of TCSC [8] [15].

2) Inductive Boost Mode:

In inductive boost mode the circulating current in the TCSC thyristor branch is bigger than the line current. In this mode, large thyristor currents result and further the capacitor voltage waveform is very much distorted from its sinusoidal shape. The peak voltage appears close to the turn on. The poor waveform and the high valve stress make the inductive boost mode less attractive for steady state operation [8].

E. ADVANTAGES OF FACTS DEVICES IN AC SYSTEMS:

Table 1: Advantage of FACTS devices in AC Systems

FACTS DEVICE	Power Flow	Voltage Control	Transient Stability	Oscillation Damping
SVC	*	***	*	**
STATCOM	*	***	**	**
TCSC	**	*	***	**
SSSC	***	*	***	**
UPFC	***	***	***	***

Influence: * Small, ** Medium, *** Strong

There are various types of FACT controllers in the market having their own advantages depending upon their individual characteristics .Here comparison study has been done on different FACT controllers to study their impact on various application, Table 1 shows the overall advantages of FACT devices in AC systems [10] [16].

IV. MODELING OF THE TCSC

The circuit is modeled using Simulink in MATLAB and ORCAD. It uses the Simulink environment, allowing user to build a model using simple click and drag procedures. It can help draw the circuit topology rapidly, and can help in the analysis the circuit and its interactions with mechanical, thermal, control, and other disciplines. This is possible because all the electrical parts of the simulation interact with the extensive Simulink modeling library. Since Simulink uses MATLAB as its computational engine, we can also use MATLAB toolboxes and Simulink block sets. Here in mat lab various blocks like voltage generator, series RLC circuit block, series RLC load branch block, current meter, voltage meter, scopes to view various signals, power meter block, thyristor block, demultiplexer bus block all are interconnected to make an open loop TCSC Simulink model which is connected in series with the single source transmission line [14].

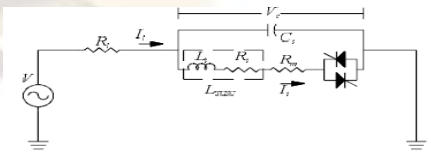


Figure 5. TCSC Circuit for simulation

For analyzing the thyristor, capacitor current and capacitor voltage, firing pulses are given to the circuit through pulse generator. To analyze the circuit in capacitive mode and inductive mode of TCSC, the pulse to be applied

are in region of Vernier capacitive which lies in 160°to 180° and 90° to 135° i.e. the delay of 8.8ms to 10ms and 5ms to 7.5ms respectively [11]. This gives the analysis of waveform of capacitor voltage, line current, thyristor current and capacitor current of TCSC as shown in fig.. It has been concluded that at if alpha=117° for positive waveform then there should be a difference or phase delay of 180° for the negative waveform then only thyristor will fire at the same interval and hence the conduction time will be equal and o/p will be symmetrical [12]. In practical applications, a lookup table method is used to describe this relationship. When the desired value of reactance X_{tcsc} is determined from the power system state, the required value of the firing angle is obtained from the table which shows the reactance step response of the TCSC to demanded change in from 1 p.u. to 2.1p.u., corresponding to a change in firing angle from to 175° to151°. It can be seen from scope output that more than 200 ms is needed for the fundamental frequency reactance to reach its new steady-state value [13].

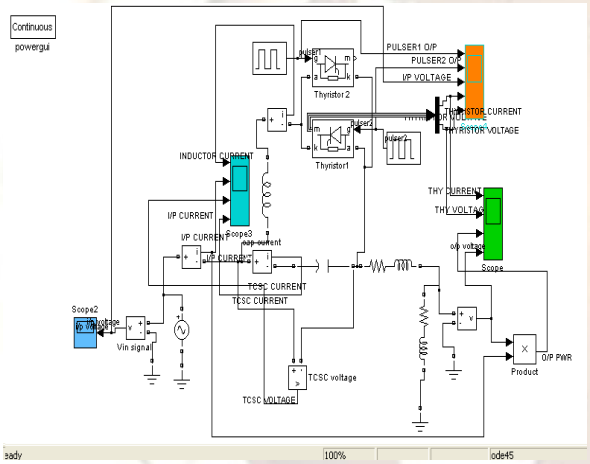


Figure 6. Simulink Model of TCSC

V. ANALYSIS OF STEADY STATE TRANSIENT BEHAVIOR THE IMPLEMENTED TCSC

Transient response of TCSC is analyzed by increasing the simulation time. The response got stable after few seconds of the firing as shown below.

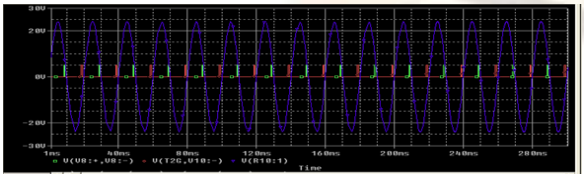


Figure 7. I/P voltage & o/p pulses of pulse generators.

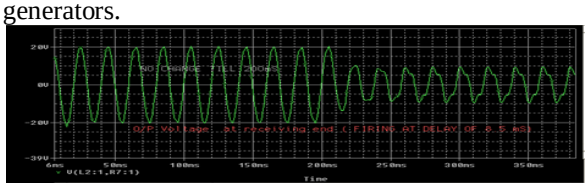


Figure 8. Output voltage a delay of 8.5ms.

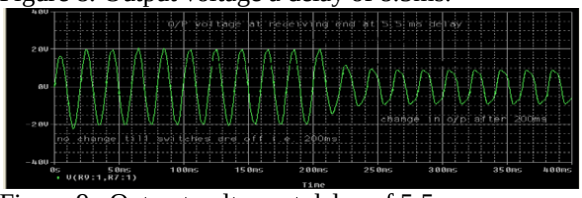


Figure 9. Out put voltage at delay of 5.5 ms

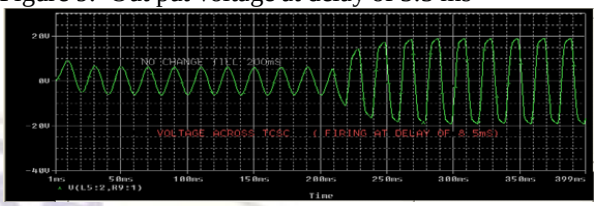


Figure 10. Voltage across TCSC at delay of 8.5ms

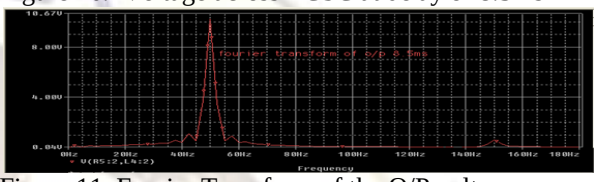


Figure 11. Fourier Transform of the O/P voltage.

Figure 12. Fourier Transform at delay=5.5 ms

A. Resultant Conclusion of TCSC modeling

Three main signals of the TCSC (line current), thyristor current (I_t) and capacitor voltage (V_c) the first one with a firing angle=16° and the 2nd one with firing angle=50°. The TCSC can be used to modify the power transfer on transmission lines because it is able to experience a fast response when the firing angle is controlled appropriately. If the transient period lasts for a long time the system may turn eventually unstable. Another noteworthy point is the fact that the signals' amplitude in the transient period remains within proper limits. This is important because if the amplitude goes too high, the components may be damaged.

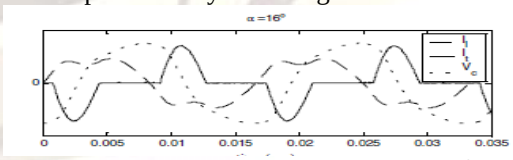


Figure13. Response when firing angle is 16°.

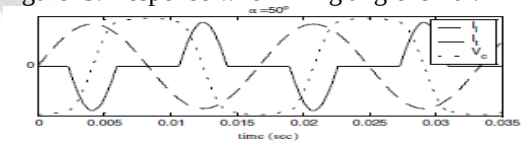


Figure14. Response when firing angle is 50°.

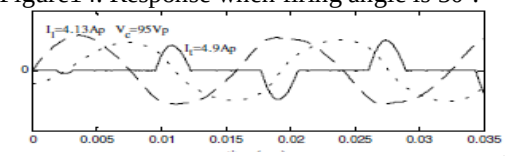


Fig.14. System response with firing angle = 42°

VI. CONCLUSION

The TCSC can be operated in capacitive and an inductive mode is rarely used in practice. The FACTS controller with its classification and performance of TCSC are included in this paper. The advantages of FACTS devices and various operating modes of TCSC are specified. This paper work can be extended in future for TCSC modeling with a number of bus system and determine the method for controlling the power flow. A linearized discrete-time model of a TCSC-compensated transmission line is presented in this paper. The model derived considers the proper characteristic of TCSC, the variation in its impedance with the firing angle of the TCR. Even though the line capacitance effects are neglected, they can be incorporated in the state equation without any major change in the model derivation, apart from an increase in the size of the matrices. The model derived is validated through digital computer simulation studies. It is shown that the eigen values of a TCSC-compensated transmission line has two complex conjugate pairs of poles whose real parts depend only on the line resistance and reactance. This model has shown the disturbances very accurately. Three different controllers are designed, based on this linearized model, which requires the measurements of local variables only. The performance of the proposed controllers under various disturbances are compared and analyzed.

Similarly in future implementation could be of a three-phase device as well as the simulation to series compensation may be by voltage sourced converters It can also be implemented with Neuro and fuzzy logic or some different strategies may now be developed to control the reactance of the TCSC through firing angle adjustments. In order to verify the performance of control strategies in actual operating conditions new techniques like instead of inductive load an original load could be implied.

REFERENCES

[1] N.G. Hingorani and L. Gyugyi "Understanding FACTS Concept and Technology of Flexible AC Transmission Systems",*IEEE.*

[2] Caudio A, Zeno & T.Faur, Canada – "Analysis of SVC and TCSC Controllers in Voltage Collappse" *IEEE Trans. Power Systems, Vol: 14, 1st Februrary 1999* pp 158-165.

[3] Preeti Singh, Mrs.Lini Mathew, Prof.S.Chatterji, NITTTR, Chandigarh "Matlab based Simulation of TCSC FACTS Controller"- *2ND National Conference on Challenges & Opportunities in Information Technology (COIT-2008) - March 2008.*

[4] Hailian Xie, Lennart Angquist, Student Royal Institute of Technology, Stockholm ON "*Synchronous Voltage Reversal control of TCSC* – Impact on SSR Conditions'.

[5] T.Venegas,C.R.Fuerte-Esquivel ON "Steady State Modeling of an Advanced Series Compensator for Power Flow Analysis of Electric Networks in Phase Co-ordinates" *IEEE Transactions on Power Delivery , Vol.16,No.4,* October 2001.

[6] Mojtaba Khederzadeh and Tarlochan Singh Sidhu, Fellow IEEE ON "Impact of TCSC on the Protection of Transmission Lines" – *IEEE Transactions on Power Delivery, Vol:21, No:1,* January 2006.

[7] R. M Mathur and R. K. Verma, Thyristor-based FACTS Controllers for *Electrical Transmission Systems, IEEE press*, Piscataway, 2002.

[8] Othman HA, Ongquist L. Analytical modeling of thyristor-controlled series capacitor for ssr studies. *Tran IEEE Power Syst* 1996; 11:1–10.

[9] Rajaraman R, Dobson I, Lasseter RH, Shern Y. Computing the damping of subsynchronous oscillations due to a thyristor-controlled series capacitor. *IEEE Trans Power Delivery* 1996; 11(2):1120–1126.

[10] H.G. Han, J.K. Park, B.H. Lee, Analysis of thyristor controlled series compensator dynamics using the state variable approach of a periodic system model, *IEEE Trans. Power Deliver*. 12 (4) (1997) 1744–1750.

[11] A.D. Del Rosso, C.A. Canizares, V.M. Dona, A study of TCSC controller design for power system stability improvement, *IEEE Trans. Power Syst*. 18 (4) (2003) 1487–1496.

[12] P. Mattavelli, G.C. Verghese, A.M. Stankovitctt, Phasor dynamics of thyristor-controlled series capacitor systems, *IEEE Trans. Power Syst.* 12 (3) (1997) 1259–1267.

[13] B.H. Li, Q.H. Wu, D.R. Turner, P.Y. Wang, X.X. Zhou, Modeling of TCSC dynamics for control and analysis of power system stability, *Int. J. Elec. Power Energy. Syst.* 22 (1) (2000) 43–49.

[14] S.Meikandasivam, Rajesh Kumar Nema and Shailendra Kumar Jain,"Behavioral study of TCSC device A MATLAB/Simulink implementation", *International Journal of Electronics and Electrical Engineering, vol. 2, no. 10, Oct. 2011*, pp. 151– 161

[15] Venu Yarlagadda, Dr. B.V Sankar Ram and Dr. K. R. M Rao, "Automatic Control of Thyristor Controlled Series Capacitor (TCSC)", *Vol. 2, Issue 3,* May-Jun 2012, pp. 444-449.

[16] Ying Xiao, Y.H. Song and Y.Z. Sun, "Power flow Control approach to power systems with embedded FACTS devices", *IEEE transaction on power Systems, vol 17, No.4, Nov. 2002*, pp. 943– 950

Author'sProfile:

Kusum Arora received M.Tech Degree and AMIE Degree from MDU, Rothak & Institution of Engrs., Kolkata, in 2010 and 2002 respectively. She is part of YMCA University of Science & Technology. She is pursuing Ph.D in 'Enhancing Power System Quality using FACTS Device'.

Dr.S.K.Agarwal received the M.Tech Degree and PhD degree in Electronics Engg. from Delhi Technical University, N. Delhi and Jamia Millia Islamia Central University, N.Delhi in 1998 and 2008 respectively. Since 1990, He has been part of YMCA University of Science & Technology Faridabad (Haryana), as Dean and Chairman in Department of Electronics Engineering. He has more than 40 publications in journals and national/international conferences of repute. His current research interests are in the field of Control System, electronics, biosensors, Analog Electronics, wireless communication and digital circuits.

Dr. Narendra Kumar received the M-Tech degree from AMU, Aligarh in 1987 and PhD in Power Systems from Univ. of Roorkee in 1996. He is part of Delhi College of engineering, New Delhi as Professor & HOD (Elect. Engg). He has more than 108 publications in journal, national/international conferences of repute. He has under taken an R&D project on 'FACTS" and has written 3 books.

DharamVir received M.Tech Degree & B.E. Degree in ECE from MDU, Rothak & Jamia Millia Islamia Central University, New Delhi in 2008, 2004 respectively. He is part of YMCA University of Science & Technology as HOS (EIC). He is pursuing PhD in 'Power aware routing in Mobile Ad hoc Networks'.

IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 6, Issue 2 (May. - Jun. 2013), PP 56-59
www.iosrjournals.org

Design of a Dual-Band Microstrip Patch Antenna for GPS,WiMAX and WLAN.

Amit Kumar[1],Sachin Kumar,Prof.P.R.Chadha[1]
[1](Dept.of ECE,Delhi Technological University,Delhi,India)

Abstract : *The A multi band microstrip patch antenna has been designed for GPS,WiMAX and WLAN applications. The proposed antenna is designed by using substrate of RT duroid having permittivity of about 2.2 and loss tangent of 1.The substrate is having thickness of 6mm at which a trapezoidal patch antenna with V slot has been introduced in this paper. The designing results like S11 parameter return loss,VSWR and field pattern is plotted successfully. The obtained result is having a two band resonance with S11 less then -10dB and VSWR less than 2.*
So a dual band trapezoidal microstrip patch antenna has been designed and all results are plotted.Simmulating software used is IE3D.

Keywords - *V-shape slot, RT duroid, Dual band, WLAN, WiMAX,*

I. INTRODUCTION

Microstrip antenna is the ideal choice for wireless an such typed application due to low profile, light weight, conformal shaping, low cost ,simplicity of manufacturing and easy integration to circuit[1].however, conventional microstrip patch antenna suffers from very narrow bandwidth, typically about 5% bandwidth with respect to the central frequency. there are numerous and well-known method to increase the bandwidth of antennas, including increase of the substrate thickness, the use of a low dielectric substrate, the use of multiple resonators,and the use of slot antenna geometry[2],[3].

Wireless local area networks (WLAN) are widely used worldwide. The 802.11a standard uses the 5-GHz band which is cleaner to support high-speed WLAN. However, the segment of frequency band used varies from one region of the world to another[6][7]. Dual frequency microstrip antennas with a single feed are required in various radar and communication systems, such as global positioning system (GPS), WiMAX,WLAN etc[1].

These communication system applications include fixed broadband local multipoint communication services, small mobile units, laptops and remote-sensing devices [2]. Also, bandwidth should be further enhanced in order to increase the information transfer rate, without sacrificing the performance[4]. There are lots of communication schemes that make use of the large operational bandwidth[8][9]. For example, the Orthogonal frequency-division multiplexing (OFDM) scheme transmit and receive signals with a number of frequency components [3]-[5]. Another example is the transmission of broadband pulse.
radiating edge.

In this paper, we design a trapezoidal patch with V-shaped antenna which works as a dual frequency. First resonance frequency f1 centered at 3.5 GHz frequency is due to its patch itself. Second resonant frequency f2 is due to V-shape slot, which is centered at 5.0 GHz.

II. ANTENNA STRUCTURE

The configuration of proposed antenna is shown in figure 1. The antenna consist of a trapezoidal microstrip patch with V-shaped slot, support on a grounded dielectric sheet of thickness h and dielectric constant ε_r. The trapezoidal patch has an upper side of length L1, base of trapezoidal patch of length L2 and height of trapezoidal patch of length W1, W2. V-shape slot has a length of L3, L4 and a width of W3, W4 which is loaded on trapezoidal patch. The feed point is located at the central line of the patch, with a distance of $d_f(x.y)$ from the bottom edge of trapezoidal patch.The dimension of trapezoidal patch with V-shape slot are tabulated in table 1.

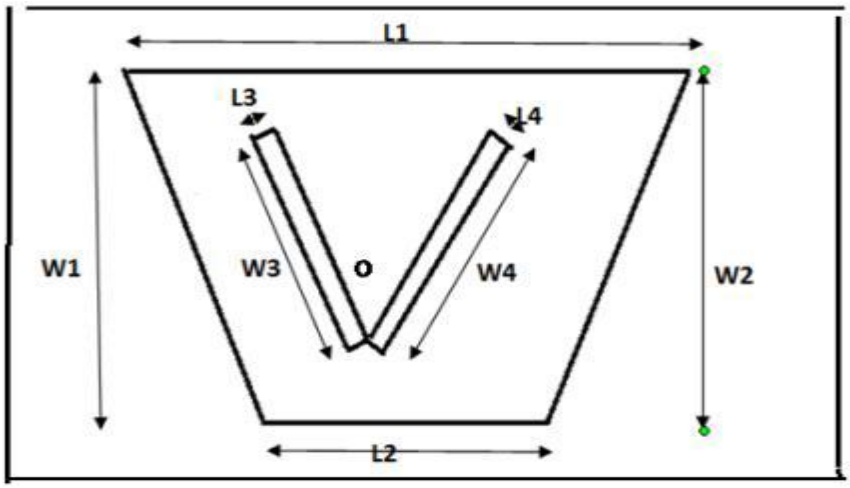


Fig.1: Geometry of proposed antenna

Table 1. Dimension of proposed antenna

S.no	Parameter	Value(mm)
1	L1	30
2	L2	26
3	W1,W2	21.04
4	L3,L4	0.5
5	W3,W4	1.5
6	ε r	2.2
7	h	6

The proposed antenna is designed and simulated by using IE3D software. All the proposed design parameters are calculated by using conventional formula for patch antenna design process.

III. DESIGN RESULTS

In this section, the simulated results of various parameters like VSWR, Return loss, input impedance and radiation characteristics of proposed antenna are presented and discussed. The simulated results are obtained using IE3D Simulator.

A. *S11 PARAMETER and VSWR*

The simulated result for the return loss less than -10dB is shown in figure 2. From simulated result we get dual band. Based on a -10 dB return loss, 4% impedance bandwidth is obtained at first resonant frequencies f1 in the frequency range (3.41-3.57) GHz and 15.6 % impedance bandwidth is obtained at second resonance frequencies in the frequency range of 4.75-5.53 GHz.

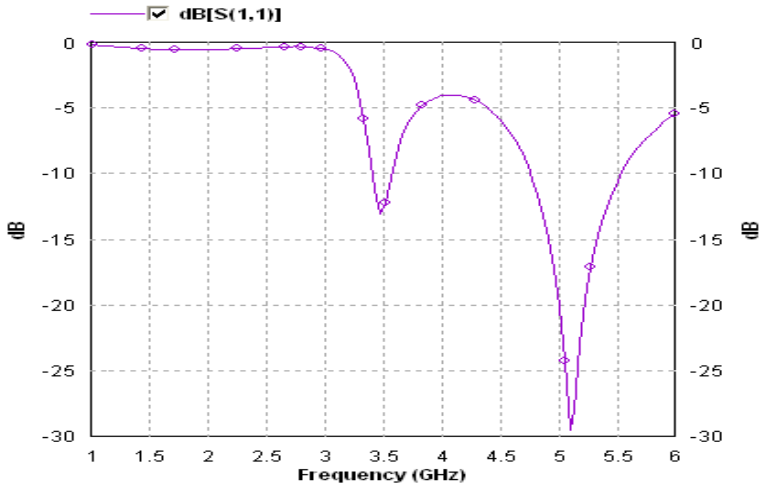


Fig. 2: Return loss

VSWR plot shows that the VSWR occur at first resonant frequency is 1.66 and second resonant frequency is 1.07. This depict that there is good impedance matching between probe-fed microstrip transmission line and the trapezoidal radiating element.

B. *Input impedance*

The simulated result for the antenna input impedance is plotted in figure 4. It is shown that the real part of the input impedance at first resonant frequency f1 oscillates around 74.83Ω with frequency while the imaginary part of the input impedance at resonant frequency oscillates around 0Ω with frequency.
At second resonant frequency f2, the real part of the input impedance at resonant frequency oscillates around 50Ω with frequency while the imaginary part of the input impedance at resonant frequency oscillates around 0Ω with frequency. Hence, from the graph it is clear that there is proper matching occur at both resonant frequencies

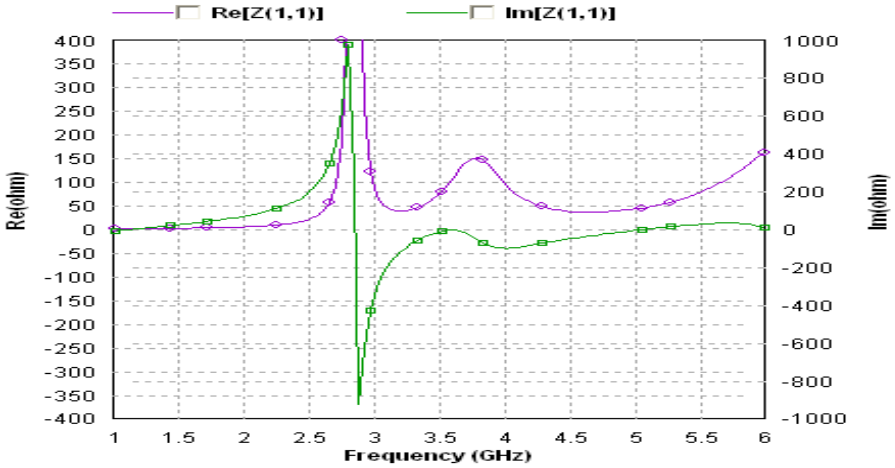


Fig. 4: Real part and imaginery part of input impedance

C. *Radiation pattern*

From figure below Shows the measured radiation pattern at first resonant frequency 3.5 GHz. it can be observed that in the ϕ=0 plane, the cross polarization is -13 db below the co polarization above the ground plane. In the ϕ=90 plane, the cross polarization is -19.3 dB below the co polarization level.

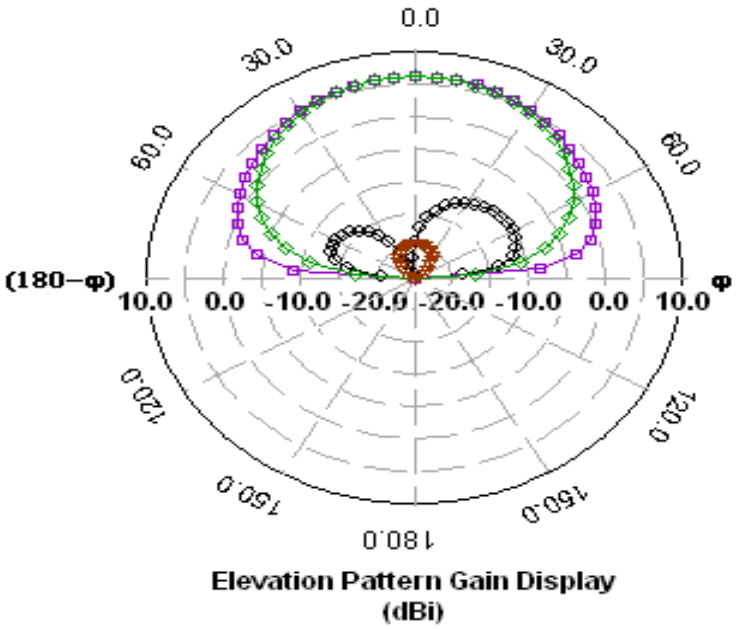


Fig. 5: Radiation pattern at 3.5 GHz

Table 2. Simulated data

Resonant frequency	F1	F2
Centre frequency	3.5 GHz	5 GHz
bandwidth	4.95	15.6%
Frequency range	(3.41-3.57)GHz	(4.75-5.53)GHz
Return loss	-12.8 dB	-29.37 dB
VSWR	1.66	1.07

IV. CONCLUSION

The dual frequency and wide-band operation of a trapezoidal patch with V-shaped slot have been studies and simulated. The proposed antenna is compact, occupies small volume and has simple structure compared to other antenna design. The antenna offer a 2:1 VSWR bandwidth of 4% from frequency range (3.41-3.57)GHz at first resonant frequency which cover 3.5 GHz band WiMAX applications. second resonant frequency cover the WLAN(5.15-5.35) band application with impedance bandwidth 0f 15.6%.the simulated return loss, VSWR, radiation pattern and gain showed well performance.

REFERENCES

[1] W.He, R.Jin, and J.Gerg, "E-shape patch with wideband & circular polarisation for millimetre wave communication," *IEEE Trans. Antenna Propag*.,vol.56,no.3, pp.893-895,2008.

[2] K.L Lau, K.M. Luk, and K.L.Lee, "Design of a circularly-polarized vertical patch antenna,"*IEEE Trans. Antenna Propag.*,vol.54, no.3, pp.1332-1335,2006.

[3] D.M Pozar & D.H Schauber, "Design of Microstrip antennas and arrays, New York*: IEEE Press*, 1995

[4] Sudhir Bhaskar & Sachin K. gupta, "Bandwidth improvement of microstrip patch antenna using H- shaped patch", *Publication in the international Journal of engineering Research and application*", vol.2, Issue 1, pp.334-338, Jan-Feb 2012.

[5] Wong KL and Hsu WS, "broadband triangular microstrip antenna with U shape slot" *Electron lett(UK)*,33(1997)2085.

[6] Sharma V, Sharma V K,Bhatnager D, Saini J.S, " Compact dual frequency wide band circular patch antenna with U-slot," *Proc of IEEE International symposium on Antenna and propoagation and USNC/URSI National Radio Science Meeting*, (IEEE,USA),1979,1.

[7] **S**. W. Lee, S. M. Park, N. Kim, S. W. Park, and S. Y. Rhee, "Design and SAR Measurement of the Trapezoidal Shape Antenna," *Progress In Electromagnetics Research C*, Vol. 26, 127-136, 2008.

[8] Amit Kumar,P.R.Chadha, Microstrip Antenna for WLAN Application Using Probe Feed, *IOSR-JECE, Volume X, Issue X (Jan. - Feb. 2013)*

[9] Prof. P.R. Chadha, Amit kumar, Ractangular Microstrip Patch Antenna Designfor WLAN Application Using Probe Feed. IJETAE, Volume 2, Issue 12, December 2012

Development of a Framework for Test Case Prioritization Using Genetic Algorithm

Ruchika Malhotra
Department of Software Engineering
Delhi Technological University
ruchikamalhotra2004@yahoo.com

Divya Tiwari
Department of Software Engineering
Delhi Technological University
divya.dec05@gmail.com

ABSTRACT

Software Testing is a time and effort consuming part of the software development life cycle. Retesting a software application during the maintenance phase, with the entire test suite and additional test cases for the modifications in the software, within budget and time, is a challenge for software testers. Test Case Prioritization is used to overcome this problem by prioritizing the test cases in order to maximize certain testing objectives like fault detection rate, statement coverage, etc. In this paper, we propose a framework for test case prioritization that emphasizes a new metric, $APBC_m$ (modified Average Percentage of Block Coverage). This metric evaluates the rate of code coverage by incorporating knowledge about the significance of blocks of code in the form of weights. We have used this metric as fitness evaluation function in a Genetic Algorithm in order to evaluate the effectiveness of a test case sequence. We have also developed a tool that implements the Genetic Algorithm in the Java language in order to compute and validate the results. Thereafter, we use the tool for test case prioritization, and compare and assess the results with those produced by the tool when APBC (Average Percentage of Block Coverage) is used as fitness function.

Categories and Subject Descriptors

D.2 [Software Engineering]: D.2.8 Metrics, D.4.8 Performance

General Terms

Measurement, Performance, Reliability, Verification.

Keywords

Regression Testing, Test Case Prioritization, Basic Block, Genetic Algorithm, Average Percentage of Block Coverage.

1. INTRODUCTION

"Software Testing is the process of executing a program with the intent of finding errors" **[1]**. ***"It is an investigation that is conducted to provide stakeholders the information about the quality of the product/service under test"*** **[2]**. The test case prioritization problem is defined as [3]:

Given a Test Suite ***'T'*** ,a set '***PT***' of all permutations of T and function '***f***' 'that maps PT to real numbers, test case prioritization technique aims to find a **T'∈ PT** such that **(∀T'') (T''∈ PT) [f(T') ≥ f(T'')].**

Empirical studies performed till now, comparing the different techniques of test case prioritization, have used either APBC (Average Percentage of Block Coverage) or APFD (Average Percentage of Fault Detection) metrics. These metrics reveal the rate at which the faults are discovered or the rate at which code coverage is achieved. Still there is a drawback of using these metrics as discussed by Elbaum et al. [4]. The APFD metric requires faults to be known prior to prioritization and treats all faults equally severe. Elbaum et al. [4] tried to improve the APFD metric by incorporating knowledge like fault severity and the cost of executing a test case. Similarly, the APBC metric considers that all blocks are equally likely to contain errors. Claes et al. [5] emphasize that certain programming constructs are more error prone than others**.** Defect data can be used to identify these programming constructs. Thus, in an application, certain blocks contribute more to faults than others. Our work aims to exploit this fact and assign weights to blocks.

The rest of the paper is organized as follows: Section 2 discusses the previous work done in the field of test case prioritization. Section 3 focuses on the proposed framework for the test case prioritization problem including details of the $APBC_m$ metric, the GA based tool and the Additional Modified Lines of Code Coverage (AMLOC) graph. It also lays down certain guidelines regarding computation of weight factors that are to be included in $\boldsymbol{APBC_m}$ metric. Section 4 shows the improvement in results after application of the modified APBC metric in the tool developed for test case prioritization using the genetic algorithm. The last section deals with the conclusions and future work that needs to be looked upon.

2. RELATED WORK

A wide range of metrics for test case prioritization have been proposed and studied. Earliest techniques revolved around coverage metrics like Statement-Total, Statement-Additional, Branch-Total, Branch-Additional, Fault Exposing Potential (FEP)-Total, FEP–Additional [3],[6-11]. Jones et al.[12] used Modified Condition/Decision Coverage and applied a greedy approach for prioritization. The main idea behind using a coverage based metric was that maximizing structural coverage of code may maximize the fault detection. The resulting prioritized test case sequences produced by various techniques were compared, for effectiveness using the APFD metric or the APBC metric. Several other machine learning techniques have also been employed in this area. Leon and Podgurski [13] have used clustering techniques to distinguish test cases associated with highly error prone regions of code from those associated with less error prone regions. They prioritized the test cases based on the density of clusters formed by the test cases. Recently Carlson et al. [14] also presented a clustering based approach to test case prioritization. They specifically applied four different metrics: code coverage, code complexity, fault history, and a combination of code complexity and fault detection ratio in order to prioritize test cases

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

within each cluster. Tonella et al. [15] used Case Based Reasoning (CBR) to prioritize the test cases. They included human knowledge of test cases for pair wise test case comparison. Yoo et al. [16] combined a human based prioritization technique with clustering technique. Kim and Porter [17] took an execution history based approach, borrowing from statistical quality control. Mirab et al. [18] exploited Bayesian networks for test case prioritization. They used information like fault proneness, code coverage, modified elements in program while providing feedback to Bayesian networks. Several model based test case prioritization techniques have also been proposed in the literature, including work of Korel et al. [19-21], Mall et al. [22]. Relevant to the work presented in this paper is the work of Roongruangsuwan et al. [23], Elbaum et al. [3] and several other cost effective Prioritization techniques [8, 24-27]. Roongruangsuwan et al. [23], in their paper on the test case prioritization technique with practical weight factors, emphasized the need to include practical weight factors. Elbaum et al. [3] emphasizes improving the APFD metric by incorporating the fact that all faults are not equally severe. The drawback of the APFD metric is that it requires faults to be known beforehand. Similar to his work, the idea presented here focuses on improving the APBC metric, taking into account that not all blocks of code are equally likely to have errors. In contrast to most of the cost effective prioritization techniques, the $APBC_m$ assigns weights to blocks of code rather than assigning weights to test cases.

3. PROPOSED FRAMEWORK FOR TEST CASE PRIORITIZATION

In this section, we propose the framework for prioritizing test cases.

3.1. The Framework

The proposed framework includes three major components- the GA based test case prioritizing tool, the Modified APBC Metric ($APBC_m$) and the Additional Modified Lines Of Code Coverage (AMLOC) graph. The initial input comprises of permutations of the test suite that serve as the initial population to the prioritization tool. The tool uses the $APBC_m$ metric to compare between two permutations and decide which candidate permutation is better and should be carried to next generation of the population. The final output of the tool is a prioritized test case sequence that maximizes the $APBC_m$. Figure 1 gives an overview of the proposed framework for test case prioritization.

3.2. Genetic Algorithm Based Tool

We develop a framework that employs a test case prioritization tool based on a Genetic Algorithm (GA), developed in the Java Language using Eclipse IDE. The tool also uses a new metric *modified APBC ($APBC_m$)*, explained in subsection 3.2.1, which is a modified form of original APBC metric. The tool uses the input data and the modified $APBC_m$ metric to produce the prioritized test case sequence and an Additional Modified Lines Of Code Coverage (AMLOC) graph for the sequence.

3.2.1. Modified APBC Metric ($APBC_m$)

In the proposed framework we use a GA based tool for prioritizing test cases and use ***modified Average Percentage of Block Coverage*** (**$APBC_m$**) for evaluating the fitness value of an individual in the population in GA. This metric is given as follows:

$$APBC_m = 1 - \frac{(w_1 * TB_1) + (w_2 * TB_2) + \ldots\ldots\ldots + (w_m * TB_m)}{n * (\sum_{i=0}^{m} w_i)} + \frac{1}{2n} \quad (1)$$

where ‘n’ is the number of test cases in the input test case set, ‘m’ is the

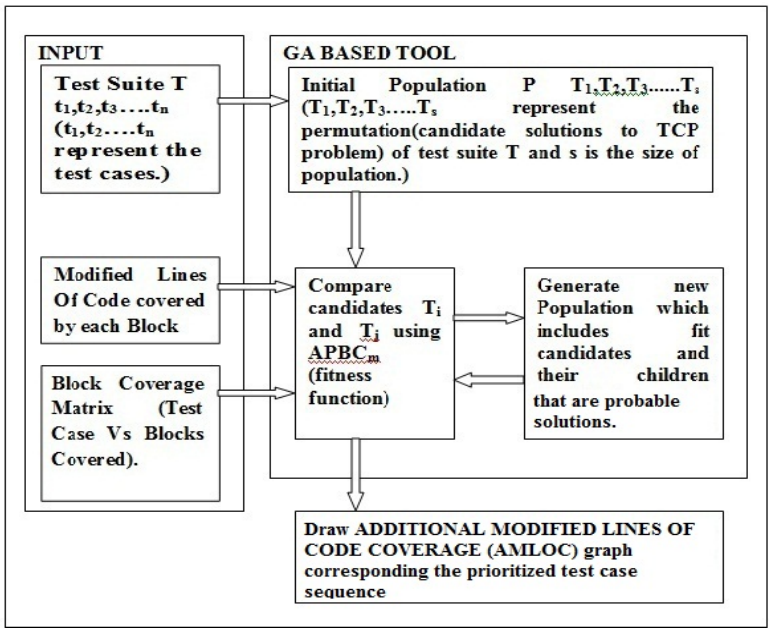


Figure 1. Framework for Test Case Prioritization

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

number of basic blocks in the program to be tested, 'TB_j' denotes the location of the test case in the test case sequence that first finds the block 'j' and 'w_i' denotes the weight of the 'i_{th}' block. APBC tries to achieve block coverage at a faster rate. However, the APBC metric has the following problems:

a. The APBC metric does not consider the practical weight factors like the significance of the blocks covered. Several factors that can be used to highlight the significance of blocks are as follows:
   i. Some blocks of code, like exception handling code, are not frequently executed. Singh et al. [31] have shown how fault proneness is correlated with object oriented metrics including design metrics. An analysis of various object oriented metrics has been done in [32].
   ii. Certain programming constructs are more error prone than others [4]. Defect data can be used to find these programming constructs. Elbaum et al. [4] have proposed a method to identify programming and design constructs that contribute more defects than expected. Ma et al. [33] have shown how analysis of program structure can be used to find important modules (e.g. methods) in a source code. Wang et al. [34] have presented a methodology using Control Flow Analysis to quantitatively analyze how the basic blocks contribute to failures.
   iii. Some blocks of code contain greater percentage of modified code. Assigning a priority to these blocks should result in faster identification of faults since modifications are most likely to contain errors.

3.2.2. Additional Modified Lines of Code Coverage (AMLOC) Graph

The GA based tool produces a prioritized test case sequence along with it's AMLOC graph. For a test case sequence TS' t_1, t_2, t_3,….t_{p-1}, t_p….t_n, the AMLOC value corresponding to a test case t_k, with respect to TS', is the ratio of the total number of unique modified lines of source code that are covered or reached by executing the test cases t_1, t_2,….t_k, where k<=n, to the total number of modified lines of source code. We use the AMLOC graph to validate the results.

4. RESULT ANALYSIS

In this section we present and validate the results produced by the framework proposed in this paper.

4.1. Experimental Study

The ***Triangle*** benchmark program has been used previously by several researchers in Software Engineering studies. Given three sides of a triangle, the program aims to classify it as a scalene, isosceles, equilateral or not a triangle. For the purpose of our experimental study, we use a simplified version of the original ***Triangle*** program presented in **[34],** translated from 'FORTRAN' into 'C'. The original set of 14 test cases was reduced to 10 test cases taking into account redundancy of test cases in terms of block coverage. We use the basic blocks identification algorithm [36] to identify the basic blocks of the source code [35]. Table 1 shows the original set of test cases.

In the code, 21 basic blocks were identified by the algorithm. The identified blocks and block coverage are given in Table 2 and Table 3 respectively. From the Table 3 it can be seen that the test cases T1, T2, T3 and T11, T13, T14 are redundant in terms of block coverage. The minimized test set includes T3, T4, T5, T6, T7, T8, T9, T10, T11, T12 and discards T1, T2, T13, T14.

Table 1. Original Set of Test Cases [35]

Input	Expected Output	Input	Expected Output
0,0,0	4 (Not a triangle)	1,2,1	4 (Not a triangle)
1,0,0	4 (Not a triangle)	2,1,1	4 (Not a triangle)
1,1,0	4 (Not a triangle)	3,2,2	2 (Isosceles triangle)
1,1,1	3 (Equilateral triangle)	3,2,1	4 (Not a triangle)
2,2,1	2 (Isosceles triangle)	4,3,2	1 (Scalene triangle)
1,1,2	4 (Not a triangle)	2,3,1	4 (Not a triangle)
2,1,2	2 (Isosceles triangle)	2,1,3	4 (Not a triangle)

Table 2. Lines of Code Comprising individual Blocks

Block	Lines Comprising the Block	Block	Lines Comprising the Block
B1	1-4	B12	21-24
B2	5-6	B13	25
B3	7-8	B14	26-27
B4	9-10	B15	28
B5	11	B16	29-30
B6	12-13	B17	31
B7	14	B18	32[-33
B8	15-16	B19	34
B9	17	B20	35-36
B10	18	B21	37-38
B11	19-20		

For the purpose of our experiment we create another modified version of ***Triangle*** by introducing the following features.

i. If the triangle is isosceles print the height of the triangle.
ii. If the triangle is scalene, print the area, circumradius and inradius of the triangle.

Table 4 shows the weight of blocks after modifying Triangle program.

The fraction of modified lines contained within each block is taken as the block weight. The idea behind doing so is that, the greater the number of modifications in a block, the greater is the error proneness of the block. In doing so, some blocks were found to have weight 0. In order to overcome this **"zero weight problem",** '1' is added to each of the values. The weights are calculated for the blocks using the modified Triangle program and are shown in table 4.

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

Table 3. Block Coverage Matrix

Test Cases	Blocks Covered	Test Cases	Blocks Covered
T1	B1,B2	T8	B1,B3,B5,B6,B7,B9, B13,B15,B17,B19,21
T2	B1,B2	T9	B1,B3,B5,B7,B8,B9, B13,B15,B17,B19,B21
T3	B1,B2	T10	B1,B3,B5,B7,B8,B9, B13,B15,B17,B19,B20
T4	B1,B3,B4,B5, B6,B7, B8,B9,B13,B14	T11	B1,B3,B5,B7, B9,B10,B11
T5	B1,B3,B4,B5, B7,B9,B13 B15,B16	T12	B1,B3,B5,B7, B9,B10,B12
T6	B1,B3,B4,B5, B7,B9,B13 B15,B17,B19, B21	T13	B1,B3,B5,B7, B9,B10,B11
T7	B1,B3,B5,B6, B7,B9, B13,B15,B17, B18	T14	B1,B3,B5,B7, B9,B10,B11

Table 4. Weight of Blocks

Block	Number of Modified Lines Of Code NMLOC	Weight NMLOC / ∑NMLOC
B1	1+1	0.0645
B2	0+1	0.032
B3	0+1	0.032
B4	0+1	0.032
B5	0+1	0.032
B6	0+1	0.032
B7	0+1	0.032
B8	0+1	0.032
B9	0+1	0.032
B10	0+1	0.032
B11	0+1	0.032
B12	6+1	0.226
B13	0+1	0.032
B14	0+1	0.032
B15	0+1	0.032
B16	1+1	0.0645
B17	0+1	0.032
B18	1+1	0.0645
B19	0+1	0.032
B20	1+1	0.0645
B21	0+1	0.032
Total	31	0.996~1

In our study, two experiments were performed. **Experiment 1** produced the prioritized test case sequence **TS1** using the data from Table 1 (after removing the redundant test cases) as input and **APBC** as fitness function in the Genetic Algorithm. **Experiment 2** produced the prioritized test case sequence **TS2** using the data from Table 1 (after removing the redundant test cases) as input and $\mathbf{APBC_m}$ as the fitness function in the Genetic Algorithm. Both experiments were performed using the Test Case Prioritization Tool based on the GA. The convergence criteria used in the tool is the maximum fitness of any individual contained in a population which was set to 0.78. We then compare the results of Experiment 1 and Experiment 2.

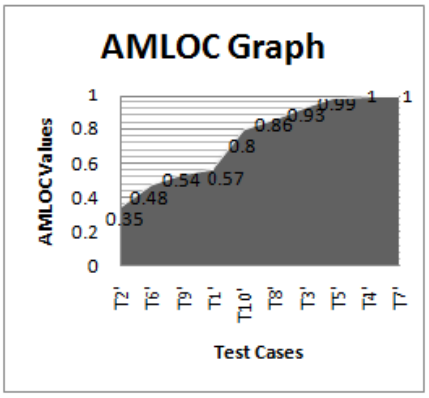


Figure 2. Less AMLOC values per test case in TS1 produced in Experiment 1

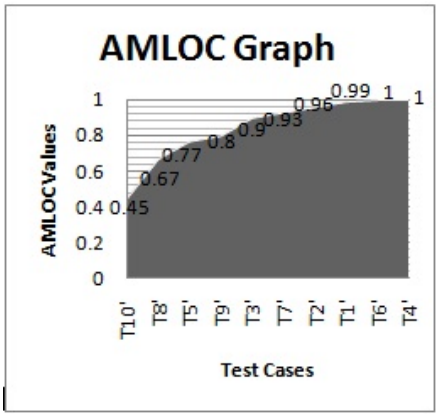


Figure 3. Greater convexity (improved coverage rate) graph using $\mathbf{APBC_m}$ in Experiment 2.

4.1. Validation Of Results

The results of prioritization, as obtained by our tool are shown in Figure 2 and Figure 3. The values calculated as per ***modified APBC and APBC*** are shown in table 5 and table 6. From the tables it is clear that using APBC for comparing the candidate solutions may produce ineffective prioritized test case sequences as output. TS2 is better as it covers the block with a greater number of modified lines first (i.e. the blocks with large weights). Hence, Modified APBC is better and is expected to provide significant help in version specific test case prioritization and help in revealing faults earlier.

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

Table 5. Results of Experiment 1

Test Case sequence / Metric	APBC
TS1(T2',T6',T9',T1',T10',T8',T3',T5',T4',T7')	0.79285

Table 6. Results of Experiment 2

Test Case sequence / Metric	Modified APBC
TS2(T10',T8',T5',T9',T3',T7',T2',T1',T6',T4')	0.80285

5. CONCLUSIONS

Test Case Prioritization is an essential task that reduces the testing effort in the maintenance phase to a considerable extent. In this paper we propose a framework for test case prioritization using a tool based on a GA, developed in the Java language. The framework also highlights the necessity and the benefits of using the new metric $APBC_m$ as fitness evaluation function in the GA. Several factors that can be used to embed the knowledge about significance of blocks in the $APBC_m$ metric have also been discussed. However, the exact computation of weights taking into account all the factors discussed, is still an open challenge. Finally, the results have been analyzed and compared with those produced when traditional APBC metric is used as fitness evaluation function in GA based tool. It was then found that $APBC_m$ metric is better and more efficient than APBC.

The approach, presented here, has its application in the areas of version specific test case prioritization but can also be generalized with slight modifications. Considering practical weight factors is a general concept that can help to improve the cost of regression testing and can also be extended to the problems of test suite minimization and regression test selection.

REFERENCES

1. Aggarwal, K.K., Singh, Y. *"Software Engineering"*. New Age International Publishers, Second ed., 2006.
2. Kaner, C., "Exploratory Testing" in *Quality Assurance Institute Worldwide Annual Software Testing Conference*, (Florida Institute of Technology Orlando, FL, 2006).
3. Chu, C., Rothermel, G., Untch, R.J. "Prioritizing test cases for regression testing" in *IEEE Transactions on Software Engineering* (October 2001), 27(10), 929–948.
4. Elbaum, S., Kanduri, S., Malsihevsky, A.G., Rothermel, G. "Selecting A Cost Effective Test Case Prioritization Technique". *Software Quality Control Journal*, 12 (3),2004, 185-210.
5. Claes, W., Magnus, C.O., Martin, H. "Understanding The Sources Of Software Defects: A Filtering Approach" in *Proceedings the 8th International Workshop on Program Comprehension, (Limerick, Ireland, 2000)*, 9-17.
6. Chu, C., Harrold, M.J., Rothermel, G., Untch, R.H. "Test case prioritization: An empirical study" in *Proceedings of International Conference on Software Maintenance*, (1999), IEEE Computer Society Press, 179–188.
7. Elbaum, S.G., Malishevsky, A.G., Rothermel, G. "Prioritizing test cases for regression testing" in *Proceedings of International Symposium on Software Testing and Analysis*, (2000), ACM Press, 102–112.
8. Elbaum, S., Gable, D., Rothermel, G. "Understanding and measuring the sources of variation in the prioritization of regression test suites" in *Proceedings of the Seventh International Software Metrics Symposium*, (2001), IEEE Computer Society Press, 169–179.
9. Elbaum, S.G., Malishevsky, A.G., Rothermel, G. "Incorporating varying test costs and fault severities into test case prioritization". In *Proceedings of the International Conference on Software Engineering*, (2001), ACM Press, 329–338.
10. Elbaum, S., Malishevsky, A.G., Rothermel, G. "Modeling the cost-benefits tradeoffs for regression testing techniques". *Proceedings of the International Conference on Software Maintenance (ICSM 2002)*, IEEE Computer Society Press, p 230–240.
11. Davia B, Elbaum S, Kallakuri P, Malishevsky A, Rothermel G, "The impact of test suite granularity on the cost effectiveness of regression testing*" in Proceedings of the 24th International Conference on Software Engineering*, (2002), ACM Press, 130–140.
12. Harrold, M.J., Jones, J.A., "Test-suite reduction and prioritization for modified condition/decision coverage" in *Proceedings of International Conference on Software Maintenance, (*2001), IEEE Computer Society Press, 92–101.
13. Leon, D., Podgurski, A., "A comparison of coverage-based and distribution-based techniques for filtering and prioritizing test cases" in *Proceedings of the IEEE International Symposium on Software Reliability Engineering*, (2003), IEEE Computer Society Press, 442–456.
14. Carlson, R., Denton, A., Do, H. "A Clustering Approach to Improving Test Case Prioritization: An Industrial Case Study" in *27th IEEE International Conference on Software Maintenance* (2011).
15. Avesani, P., Susi, A., Tonella, P. "Using the case-based ranking methodology for test case prioritization" in *Proceedings of the 22nd International Conference on Software Maintenance*, (2006), IEEE Computer Society Press, 123–133.
16. Harman, M., Susi, A., Tonella, P., Yoo, S. "Clustering test cases to achieve effective & scalable prioritization incorporating expert knowledge" in *Proceedings of International Symposium on Software Testing and Analysis*, (2009), ACM Press, 201–211.
17. Kim, J.M., Porter, A. "A history-based test prioritization technique for regression testing in resource constrained environments" in *Proceedings of the 24th International Conference on Software Engineering*, (2002), ACM Press, 2002, 119–129.
18. Mirarab, S., Tahvildari, L. "A prioritization approach for software test cases based on bayesian networks" in *Proceedings of the 10th International Conference on Fundamental Approaches to Software Engineering*, (2007), Springer–Verlag, 276–290.

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

19. Harman, M., Korel, B., Tahat, L. “Test prioritization using system models” in *Proceedings of the 21st IEEE International Conference on Software Maintenance*, (2005), 559–568.
20. Korel, B., Koutsogiannakis, G., Tahat, L.H. “Model-based test prioritization heuristic methods and their evaluation” in *Proceedings of the 3rd international workshop on Advances in Model-based Testing*, (2007), ACM Press, 34–43.
21. Korel, B., Koutsogiannakis, G., Tahat, L. “Application of system models in regression test suite prioritization” in *Proceedings of IEEE International Conference on Software Maintenance*, (2008), IEEE Computer Society Press, 247–256.
22. Mall, R., Panigrahi, C.R. “Model Based Regression Test Case Prioritization”. *ACM SIGSOFT Software Engineering Notes*, 35 (6),2010, 1-7
23. Daengdej, J., Roongruangsuwan, S. “A Test Case Prioritization Method with Practical Weight Factors”, *Journal Of Software Engineering*, 4(3),2010, 193-214.
24. Harman, M., Yoo, S. “Efficient multi-objective test case selection” in *Proceedings of International Symposium on Software Testing and Analysis*, (2007), ACM Press, 140–150.
25. Do, H., Mirarab, S.M., Rothermel, G., Tahvildari, L. “An empirical study of the effect of time constraints on the cost-benefits of regression testing” in *Proceedings of the 16th ACM SIGSOFT International Symposium on Foundations of Software Engineering*,(2008), ACM Press, 71–82.
26. Kapfhammer, G.M., Roos, R.S., Soffa, M.L., Walcott, K.R. “Time aware test suite prioritization” in *Proceedings of the International Symposium on Software Testing and Analysis*, (2006), ACM Press, 1–12.
27. Qu, B., Zhang, X. “An Improved Metric for Test Case Prioritization” in *Eighth Web Information Systems and Applications Conference*, (2011), IEEE.
28. Johnson,S.M. *"Generation of permutations by adjacent transposition"*, Mathematics of Computation 17, 282–285, doi:10.1090/S0025-5718-1963-0159764-2.
29. Trotter,H.F. "Algorithm 115: Perm", *Communications of the ACM* 5 (8), 1962, 434–435.
30. Levitin, A. *“Introduction to The Design & Analysis of Algorithms”*, Addison Wesley, 2003.
31. Kaur, A., Malhotra, R., Singh, Y. “Software Fault Proneness Prediction Using Support Vector Machines” in *Proceedings of the World Congress on Engineering*, (London, U.K.,2009),1 – 3.
32. Aggarwal, K. K., Kaur, A., Malhotra, R., Singh, Y. "Analysis of object-oriented metrics." in *International Workshop on Software Measurement* 2005.
33. Ma, Z., Zhao, J. “Test Case Prioritization based on the Analysis Of Program Structure” in *15th Asia-Pacific Software Engineering Conference*, (2008, IEEE, 471-478.
34. Wang, L., Yin, X., Zhao, Lei. “Context Aware Fault Localization via Control Flow Analysis”. *Journal Of Software, 6(10)*, 2011.
35. Harman, M., Jia, Y., Langdon, W.B. “Efficient multi-objective higher order mutation testing with genetic programming”. *Journal Of Systems and Software*, 83(12), 2416-2430.
36. Allen, F.E. “Control Flow Analysis” in *Proceedings of Symposium on Compiler Optimization*, (1970), 5(7), 1-19.

DOI: 10.1145/2464526.2464536 http://doi.acm.org/10.1145/2464526.2464536

Astrophys Space Sci
DOI 10.1007/s10509-013-1436-z

ORIGINAL ARTICLE

Functional form of $f(R)$ with power-law expansion in anisotropic model

Vijay Singh · C.P. Singh

Received: 30 November 2012 / Accepted: 26 March 2013

Abstract In this paper, we study an anisotropic Bianchi-I space-time model in $f(R)$ theory of gravity in the presence of perfect fluid as a matter contains. The aim of this paper is to find the functional form of $f(R)$ from the field equations and hence the solution of various cosmological parameters. We assume that the deceleration parameter to be a constant, and the shear scalar proportional to the expansion scalar to obtain the power-law form of the scale factors. We find that the model describes the decelerated phases of the universe under the choice of certain constraints on the parameters. The model does not show the acceleration expansion and also transition from past deceleration to present accelerating epoch. We discuss the stability of the functional form of $f(R)$ and find that it is completely stable for describing the decelerating phase of the universe.

Keywords $f(R)$ gravity · Anisotropic models

1 Introduction

The recent developments in cosmology with the observations such as Ia supernova (Riess et al. 1998, 1999; Perlmutter et al. 1999; Tonry et al. 2003), cosmic microwave background anisotropy (Spergel et al. 2003), large scale structure (Tegmark et al. 2004; Seljak et al. 2005; Percival et al. 2007; Kamatsu et al. 2009), baryon oscillation (Eisenstein et al. 2005) and weak lens (Jain and Taylor 2003) have led to the conclusion that the universe is accelerating in the current epoch. It has been observed that a fluid known as dark energy (DE) with large negative pressure is responsible for this acceleration. Many DE models have been proposed to explain the cosmic accelerated expansion (Copeland et al. 2006). A cosmological constant Λ responsible for acceleration is the simplest candidate of DE (Sahni and Starobinsky 2000; Padmanabhan 2003), which so far best fits with the observational data. However, the observed value of the cosmological constant is much smaller (10^{120} orders of smaller magnitude) than any other energy (vacuum energy) predicted by quantum physics (Peebles and Ratra 2003; Carroll 2001). This problem is known as fine-tuning problem in cosmology. Therefore, there is no guiding principle for construction of a promising model of cosmological constant.

The second alternative of cosmological constant, which is not stable, is a minimally coupled scalar field ϕ, usually called quintessence. These scalar fields may be responsible for a stage of accelerated expansion (Steinhardt et al. 1999; Zlatev et al. 1999). A further interesting possibility is provided by non-minimally coupled scalar field (Amendola 1999; Chiba 2003a). It is also mentioned that energy conditions are violated in all these kind of scalar fields. Also, these models end with a finite future singularity known as Big Rip. Modification of the gravity theory is an alternative approach, for example, $f(R)$ model (Capozziello et al. 2003; Nojiri and Odintsov 2003; Carroll et al. 2004). The $f(R)$ model is a modified gravity model, constructed by replacing the gravitational Lagrangian with a general function of the Ricci scalar R.

The $f(R)$ gravity provides a very natural unification of the early-time inflation and late-time acceleration. It describes the transition from deceleration to acceleration in the evolution of the universe (Nojiri and Odintsov 2007a,

V. Singh · C.P. Singh (✉)
Department of Applied Mathematics, Delhi Technological University (Formerly Delhi College of Engineering), Bawana Road, Delhi 110 042, India
e-mail: cpsphd@rediffmail.com

V. Singh
e-mail: gtrcosmo@gmail.com

Published online: 14 May 2013

2008). Over the past few years, Many works are available in literature (Capozziello and Francaviglia 2008; Abdalla et al. 2005; Nojiri and Odintsov 2007b; Bertolami and Pármos 2008; Harko 2008) addressing the well-known issues of stability (Dolgov and Kawasaki 2003), singularity problem (Frolov 2008), solar system test (Chiba 2003b), etc. The general scheme for modified gravity reconstruction from any realistic FRW cosmology have been discussed by Nojiri and Odintsov (2006). It seems that $f(R)$ gravity models pass all known observational local test currently (Elizalde et al. 2010, 2011; Nojiri and Odintsov 2011).

Almost all of these considerations are mainly investigated in a spatially flat homogeneous and isotropic universe described by Friedmann- Robertson-Walker (FRW) metric. The theoretical studies and observational data, which support the existence of anisotropic phase, lead to consider the model of universe with anisotropic background. Since, the universe is almost isotropic at large scale, the studying of the possible effects of anisotropic universe in the early time makes the Bianchi-I type (BI) model as a prime alternative. Many authors have tried to find analytical solutions for the known functional form $f(R)$. For examples, Barrow and Clifton (2006) have obtained exact cosmological solutions for scale invariant theories which generalize Einstein's GR to a theory derived from the Lagrangian $R^{1+\delta}$. The solutions were expanding universe of Kasner form and exist for $-\frac{1}{2} < \delta < \frac{1}{4}$. However, these solutions were obtained for vacuum universe. Sharif and Shamir (2009, 2010a, 2010b), Sharif and Zubair (2010a), Shamir (2010), Sharif and Kausar (2011a, 2011b, 2011c), and Aktaş et al. (2012) have studied anisotropic models in $f(R)$ theory. Recently, Yilmaz et al. (2012) have discussed quark and strange quark matter in $f(R)$ gravity for Bianchi I and V space time models. The ordering of this approach can also be reversed. Namely, for a known scale factor, one may construct functional form of $f(R)$ which yields such scale factors as solutions (Nojiri and Odintsov 2007c; Capozziello and Francaviglia 2008; Nojiri et al. 2009).

In this paper, we are interested to find a functional form of $f(R)$ for a known scale factor in anisotropic locally-rotationally-symmetric (LRS) Bianchi I model with perfect fluid as a source of matter. A functional form of $f(R)$ is obtained from the field equations by assuming the constant deceleration parameter and the shear scalar proportional to the expansion scalar. We find that the model describes the decelerated phase of the universe under the choice of certain constraints on the parameter. We discuss the stability of $f(R)$ theory for the decelerated phase of the solution and it is found that it is completely stable for the defined constraint.

The paper is organized as follows. In Sect. 2, we present the gravitational action of $f(R)$ gravity and the corresponding field equations. Section 3 provides the solution of the field equations. In Sect. 4 we discuss the stability condition for Bianchi model in $f(R)$ gravity. In the last Sect. 5, summary of the finding is given.

2 Field equations for $[f(R) + \mathcal{L}_m]$ gravity

The gravitational action for $f(R)$ theory of gravity coupled with matter fluid in the units $16\pi G = 1$ and $c = 1$, takes the following form (Nojiri and Odintsov 2007c; Capozziello and Francaviglia 2008)

$$I = \int d^4x\sqrt{-g}\big[f(R) + \mathcal{L}_m\big], \tag{1}$$

where R is the Ricci scalar and $\mathcal{L}_m$ corresponds to the matter Lagrangian.

The field equations are obtained by varying the action (1) with respect to metric tensor $g_{\mu\nu}$

$$F(R)R_{\mu\nu} - \frac{1}{2}f(R)g_{\mu\nu} - \nabla_\mu\nabla_\nu F(R) + g_{\mu\nu}\Box F(R) = T_{\mu\nu}, \tag{2}$$

where $F(R) = f'(R)$ and $T_{\mu\nu}$ is the energy momentum tensor. The other symbols have their usual meanings.

The energy momentum tensor for a perfect fluid is given as

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - pg_{\mu\nu}, \tag{3}$$

where ρ is the energy density and p is the thermodynamical pressure of the fluid. u_μ is the four velocity of the fluid such that $u_\mu u^\nu = 1$ and in comoving coordinates, $u^\mu = \delta^\mu_0$.

We consider a homogeneous and anisotropic Locally-rotationally-symmetric (LRS) Bianchi type-I line element whose metric is given by

$$ds^2 = dt^2 - A^2dx^2 - B^2\big(dy^2 + dz^2\big), \tag{4}$$

where the metric coefficients A and B are the scale factors in an anisotropic background and are functions of cosmic time t only.

The average scale factor is defined as

$$a = \big(AB^2\big)^{\frac{1}{3}}. \tag{5}$$

The rate of the expansion along x-, y-, and z-axes can be defined as,

$$H_x = \frac{\dot{A}}{A}, \qquad H_y = H_z = \frac{\dot{B}}{B}, \tag{6}$$

where a dot denotes ordinary derivative with respect to cosmic time t. The average Hubble parameter (average expansion rate), which is the generalization of the Hubble parameter in an isotropic case, H is given as

$$H = \frac{\dot{a}}{a} = \frac{1}{3}\left(\frac{\dot{A}}{A} + 2\frac{\dot{B}}{B}\right). \tag{7}$$

The expansion scalar, θ and the shear scalar, σ^2 are respectively defined as

$$\theta = u^i_{;i} = 3H = \frac{\dot{A}}{A} + 2\frac{\dot{B}}{B}, \tag{8}$$

$$\sigma^2 = \frac{1}{2}\sigma_{ij}\sigma^{ij} = \frac{1}{3}\left(\frac{\dot{A}}{A} - \frac{\dot{B}}{B}\right)^2, \tag{9}$$

where

$$\sigma_1^1 = \frac{2}{3}\left(\frac{\dot{A}}{A} - \frac{\dot{B}}{B}\right),$$
$$\sigma_2^2 = \sigma_3^3 = -\frac{1}{3}\left(\frac{\dot{A}}{A} - \frac{\dot{B}}{B}\right), \qquad \sigma_4^4 = 0. \tag{10}$$

The scalar curvature for the metric (4) is given by

$$R = -2\left(\frac{\ddot{A}}{A} + 2\frac{\ddot{B}}{B} + 2\frac{\dot{A}\dot{B}}{AB} + \frac{\dot{B}^2}{B^2}\right). \tag{11}$$

Using comoving coordinates, the field Eqs. (2) for the metric (4) and energy-momentum tensor (3) yield the following system of equations.

$$\left(\frac{\dot{A}}{A} + 2\frac{\dot{B}}{B}\right)\dot{F} - \left(\frac{\ddot{A}}{A} + 2\frac{\ddot{B}}{B}\right)F - \frac{1}{2}f = \rho, \tag{12}$$

$$\ddot{F} + 2\frac{\dot{B}}{B}\dot{F} - \left(\frac{\ddot{A}}{A} + 2\frac{\dot{A}\dot{B}}{AB}\right)F - \frac{1}{2}f = -p, \tag{13}$$

$$\ddot{F} + \left(\frac{\dot{A}}{A} + \frac{\dot{B}}{B}\right)\dot{F} - \left(\frac{\ddot{B}}{B} + \frac{\dot{A}\dot{B}}{AB} + \frac{\dot{B}^2}{B^2}\right)F - \frac{1}{2}f = -p. \tag{14}$$

From (13) and (14), we get

$$\frac{\dot{F}}{F} = -\frac{(\frac{\ddot{A}}{A} - \frac{\ddot{B}}{B} + \frac{\dot{A}\dot{B}}{AB} - \frac{\dot{B}^2}{B^2})}{(\frac{\dot{A}}{A} - \frac{\dot{B}}{B})}. \tag{15}$$

On integration of (15), we obtain

$$F = \frac{F_0}{B(\dot{A}B - \dot{B}A)}, \quad \text{provided } A \neq B, \tag{16}$$

where F_0 is a constant of integration. Equation (16) represents the most general form of $F(R)$ in terms of directional scale factors in $f(R)$ gravity for the anisotropic LRS Bianchi-I space-time.

3 Solution of the field equations

Many authors have tried to find analytical solutions from the known functional form $f(R)$ (Barrow and Clifton 2006). The ordering of this approach can also be reversed. Namely, for a known scale factor, one may construct functional form of $f(R)$ which yields such scale factor as a solution (Nojiri and Odintsov 2007c; Capozziello and Francaviglia 2008; Nojiri et al. 2009). As we observe that the solution of F(R) in (16) can be found only if the scale factors are known. In this paper our aim is to find a general form of $f(R)$ for a known scale factor and study the stability of functional form of $f(R)$ in anisotropic model to describe the decelerated and accelerated phases of the universe.

For any physically relevant model, the Hubble parameter and deceleration parameter are the most important observational quantities in cosmology. Berman (1983), and Berman and Gomide (1988) proposed a law of variation for Hubble parameter in FRW model that yields a constant value of deceleration parameter and a power-law and exponential forms of the scale factor. In recent papers (see Refs. Singh and Kumar 2006; Singh et al. 2008; Singh 2009a, 2009b; Singh and Beesham 2010; Sharif and Zubair 2010b, 2012a) have generalized this assumption in anisotropic model. According to the assumption let us take the deceleration parameter as a constant, that is,

$$q = -\frac{a\ddot{a}}{\dot{a}^2} = n - 1, \tag{17}$$

where $n(\geq 0)$ is a constant. In the present anisotropic model, the assumption (17) yields

$$a = \left(AB^2\right)^{\frac{1}{3}} = (ct + d)^{\frac{1}{n}}, \quad n \neq 0, \tag{18}$$

where c and d are positive constants of integration. For a power-law expansion (18), we must have $n > 0$.

In view of anisotropy of the space-time, we assume that shear scalar (σ) is proportional to the expansion scalar (θ). This leads to a relation between the metric coefficients, i.e.,

$$A = B^k, \tag{19}$$

where $k > 1$ is a constant (Collins et al. 1980). For sake of simplicity, we have taken the integration constant as a unity.

Using (18) and (19), we get the metric coefficients as

$$A = (ct + d)^{\frac{3k}{n(k+2)}}, \tag{20}$$

$$B = (ct + d)^{\frac{3}{n(k+2)}}. \tag{21}$$

From (11), (20) and (21), the expression for the Ricci scalar becomes

$$R(t) = -2\left[\frac{9\{k(k+2)+3\} - 3n(k+2)^2}{n^2(k+2)^2}\right]t^{-2}$$
$$= -2\alpha t^{-2}, \tag{22}$$

where $\alpha = \frac{9\{k(k+2)+3\}-3n(k+2)^2}{n^2(k+2)^2}$.

Using the above background solutions into (16), we find $f'(R)$ in terms of R as

$$F(R) = f'(R) = \frac{2^{\frac{n-3}{2n}}F_0(k+2)}{3(k-1)}\left[n\sqrt{-\frac{\alpha}{R}}\right]^{\frac{n-3}{n}}. \tag{23}$$

We observe that for a real valued solution of $f(R)$, R and α must be of opposite sign.

On integration of (23), we get

$$f(R)=\frac{2^{\frac{3(n-1)}{2n}}F_0(k+2)nR}{3(k-1)(n+3)}\left[n\sqrt{-\frac{\alpha}{R}}\right]^{\frac{n-3}{n}}+f_0, \tag{24}$$

where f_0 is a constant of integration. On imposing $f(0)=0$, we find $f_0=0$, therefore, Eq. (24) gives the required form of the function $f(R)$ as

$$f(R)=\frac{2^{\frac{3(n-1)}{2n}}F_0(k+2)nR}{3(k-1)(n+3)}\left[n\sqrt{-\frac{\alpha}{R}}\right]^{\frac{n-3}{n}}, \tag{25}$$

which is basically of the form $f(R)\propto R^{\lambda}$, where $\lambda=\frac{n+3}{2n}>0$ as $n>0$. For $n=1$, $f(R)\propto R^2$ and for $n=3$, the model reduces to general relativity form, i.e., $f(R)\propto R$. It is also clear that the power of R, i.e., λ contains only n and is independent of k.

The energy density and pressure in terms of cosmic time t are given as

$$\rho=6\left[\frac{n(k+2)^2-3\{k(k+2)+3\}}{(k+2)(k-1)(n+3)}\right]\frac{1}{(nt)^{\frac{n+3}{n}}}, \tag{26}$$

$$p=2\left[\frac{n(k+2)^2-3\{k(k+2)+3\}}{(k+2)(k-1)(n+3)}\right]\frac{n}{(nt)^{\frac{n+3}{n}}}. \tag{27}$$

For reality of the viable model, the energy density must be positive and therefore we must take $n(k+2)^2-3\{k(k+2)+3\}>0$, as $k>1$. Equations (26) and (27) give

$$\rho+p=2\left[\frac{n(k+2)^2-3\{k(k+2)+3\}}{(k+2)(k-1)}\right]\frac{1}{(nt)^{\frac{n+3}{n}}}, \tag{28}$$

which shows that null energy condition (NEC) is satisfied for $k>1$. The energy density and pressure decrease with time and tend to zero for large t. The equation of state parameter ω, which is defined as $\omega=\frac{p}{\rho}$, is given by

$$\omega=\frac{n}{3}, \tag{29}$$

which is positive through out the evolution. This shows that the functional form of $f(R)$ in (25) with the metric coefficients (20) and (21) describes the decelerated phases of the universe. From (17) and (29), we have the following linear relation in q and ω

$$q=3\omega-1. \tag{30}$$

which also shows that the model decelerates.

We now find the following constraints under which the model decelerates keeping in view of the positivity of energy density

$$2<n<3 \quad \text{and} \quad 1<k<\frac{3-2n}{n-3}+3\sqrt{\frac{n-2}{(n-3)^2}} \tag{31}$$

or,

$$n\geq 3 \quad \text{and} \quad k>1. \tag{32}$$

Therefore, we conclude that the $f(R)$ model obtained here in (25) favors the decelerating phase of the universe under above constraints (31) and (32), which is not in according to the present day scenario of the accelerating universe. Also, the model does not show transition from decelerated to accelerated phase of the universe. In literature it can be seen that the $f(R)$ models may imply accelerated expansion and also transition from past deceleration to present acceleration epoch. It may be noted that the anisotropic Bianchi models represent cosmos in its early stages of the evolution of the universe. In a paper (Sharif and Zubair 2012b), it is shown that some models can explain the evolutionary paradigm which would result in deceleration phase. Therefore, this functional form of $f(R)$ can also explain some of the physical properties of the evolution of the decelerating universe.

4 Stability analysis

In Sect. 3, we have obtained a suitable functional form of $f(R)$, which describes the decelerating phase of the universe under some defined constraints. In this section, we study the stability of a foresaid form of $f(R)$. An acceptable cosmological model in $f(R)$ theory is considered to be viable if it satisfies the following stability conditions

$$f'(R)>0, \tag{33}$$

and

$$f''(R)>0. \tag{34}$$

The conditions for the cosmological viability of $f(R)$ models have been derived in Ref. (Amendola et al. 2007). Among those conditions the requirement of the above two are particularly important to give rise to a saddle matter era followed by a late time cosmic acceleration. The cosmologically viable $f(R)$ models need to be close to the Λ-Cold Dark Matter (ΛCDM) model in the deep matter era, but the deviation from it becomes important around the late stage of the matter era. Several examples of such viable models were presented in Li and Barrow (2007), Amendola and Tsujikawa (2008).

The stability condition (33) always holds as $k>1$. To obtain the constraints which satisfy (34), we differentiate (23) with respect to R, to get

$$f''(R)=\frac{2^{-\frac{(n+3)}{2n}}F_0(k+2)n(n-3)\alpha}{3(k-1)R^2}\left[n\sqrt{-\frac{\alpha}{R}}\right]^{-\frac{n+3}{n}}, \tag{35}$$

which gives

$$n>3, \qquad k>1. \tag{36}$$

Thus, we find that out of two constraints described in (31) and (32), the constraint in Eq. (32) only favors the stability of the solution for the decelerated phases of the universe.

5 Conclusion

In this paper, we have studied $f(R)$ theory of gravity in anisotropic LRS Bianchi-I space-time model. We have assumed the constant deceleration parameter, and a proportionality relation between shear scalar and scalar expansion to obtain exact form of $f(R)$ and then the solution of the various cosmological parameters. We have found that the value of EoS parameter is constant, i.e., $\omega = \frac{n}{3}$, which remains positive as $n > 0$ throughout the evolution of the universe. This allows to describe the decelerated phases of the universe. We have obtained a linear relation between the deceleration parameter, q and EoS parameter, ω as $q = 3\omega - 1$, which also describes the decelerated epoch. In Sect. 3 we have obtained some specific constraints on parameters keeping in view of positivity of the energy density under which the model exhibits decelerated universe. The model does not show transition from decelerated to accelerated phase of the universe.

We have also analyzed the stability of the functional form of $f(R)$ and found that it is completely stable to describe the decelerated phases of the universe. We have not obtained any suitable constraints to describe the accelerated expansion and also the transition from past deceleration to present accelerating epoch of the universe in this anisotropic model. It is well known that the anisotropic models represent cosmos in its early stages of the evolution of the universe and some models can explain the evolutionary paradigm which would result in deceleration phase. It may be noted that even though the $f(R)$ gravity describes an early-time inflation and late-time acceleration, this results shows that $f(R)$ gravity theory is also suitable to describe the decelerated phase of the universe in anisotropic models which is stable. The approach introduced is simple and much more universal. We hope that this will make it useful in future applications of $f(R)$ theory of gravity in anisotropic models.

Acknowledgements The authors would like to thanks the anonymous referees for very valuable comments on this work. The authors also express their sincere thanks to Prof. Sami, Center for Theoretical Physics, JMI, India for useful discussion.

References

Abdalla, M.C.B., Nojiri, S., Odintsov, S.D.: Class. Quantum Gravity **22**, L35 (2005)
Aktaş, C., Aygün, S., Yilmaz, İ.: Phys. Lett. B **707**, 237 (2012)
Amendola, L.: Phys. Rev. D **60**, 043501 (1999)
Amendola, L., Tsujikawa, S.: Phys. Lett. B **660**, 125 (2008)
Amendola, L., Gannouji, R., Polarski, D., Tsujikawa, S.: Phys. Rev. D **75**, 083504 (2007)
Barrow, J.D., Clifton, T.: Class. Quantum Gravity **23**, L1 (2006)
Berman, M.S.: Nuovo Cimento B **74**, 182 (1983)
Berman, M.S., Gomide, F.M.: Gen. Relativ. Gravit. **20**, 191 (1988)
Bertolami, O., Pármos, J.P.: Class. Quantum Gravity **25**, 245017 (2008)
Capozziello, S., Francaviglia, M.: Gen. Relativ. Gravit. **40**, 357 (2008)
Capozziello, S., Carloni, S., Troisi, A.: Recent Res. Dev. Astron. Astrophys. **1**, 625 (2003)
Carroll, S.M.: Living Rev. Relativ. **4**, 1 (2001)
Carroll, S.M., Duvvuri, V., Trodden, M., Turner, M.S.: Phys. Rev. D **70**, 043528 (2004)
Chiba, T.: Phys. Rev. D **60**, 083508 (2003a)
Chiba, T.: Phys. Lett. B **575**, 1 (2003b)
Collins, C.B., Glass, e.N., Wilkinson, D.A.: Gen. Relativ. Gravit. **12**, 805 (1980)
Copeland, E.J., Sami, M., Tsujikawa, S.: Int. J. Mod. Phys. D **15**, 1753 (2006)
Dolgov, A.D., Kawasaki, M.: Phys. Lett. B **573**, 1 (2003)
Eisenstein, D.J., et al.: Astrophys. J. **633**, 560 (2005)
Elizalde, A., Nojiri, S., Odintsov, S.D., Gomez, D.S.: Eur. Phys. J. C **70**, 351 (2010)
Elizalde, A., Nojiri, S., Odintsov, S.D., Sebastiani, L., Zerbini, S.: Phys. Rev. D **83**, 086006 (2011)
Frolov, A.V.: Phys. Rev. Lett. **101**, 061103 (2008)
Harko, T.: Phys. Lett. B **669**, 376 (2008)
Jain, B., Taylor, A.: Phys. Rev. Lett. **91**, 141302 (2003)
Kamatsu, E., et al.: Astrophys. J. Suppl. Ser. **180**, 330 (2009)
Li, B., Barrow, J.D.: Phys. Rev. D **75**, 084010 (2007)
Nojiri, S., Odintsov, S.D.: Phys. Rev. D **68**, 123512 (2003)
Nojiri, S., Odintsov, S.D.: Phys. Rev. D **74**, 086005 (2006)
Nojiri, S., Odintsov, S.D.: Phys. Lett. B **657**, 238 (2007a)
Nojiri, S., Odintsov, S.D.: Phys. Lett. B **646**, 105 (2007b)
Nojiri, S., Odintsov, S.D.: Int. J. Geom. Methods Mod. Phys. **4**, 115 (2007c)
Nojiri, S., Odintsov, S.D.: Phys. Rev. D **77**, 026007 (2008)
Nojiri, S., Odintsov, S.D.: Phys. Rep. **505**, 59 (2011)
Nojiri, S., Odintsov, S.D., Saez-Gomez, D.: Phys. Lett. B **681**, 74 (2009)
Padmanabhan, T.: Phys. Rep. **380**, 235 (2003)
Peebles, P.J.E., Ratra, B.: Rev. Mod. Phys. **75**, 559 (2003)
Percival, W.J., et al.: Mon. Not. R. Astron. Soc. **381**, 1053 (2007)
Perlmutter, S., et al.: Astrophys. J. **517**, 565 (1999)
Riess, A.G., et al.: Astrophys. J. **116**, 1009 (1998)
Riess, A.G., et al.: Astrophys. J. **117**, 707 (1999)
Sahni, V., Starobinsky, A.A.: Int. J. Mod. Phys. D **9**, 373 (2000)
Seljak, U., et al.: Phys. Rev. D **71**, 103515 (2005)
Shamir, M.F.: Astrophys. Space Sci. **330**, 183 (2010)
Sharif, M., Kausar, H.R.: Astrophys. Space Sci. **332**, 463 (2011a)
Sharif, M., Kausar, H.R.: J. Phys. Soc. Jpn. **80**, 044004 (2011b)
Sharif, M., Kausar, H.R.: Phys. Lett. B **697**, 1 (2011c)
Sharif, M., Shamir, M.F.: Class. Quantum Gravity **26**, 235020 (2009)
Sharif, M., Shamir, M.F.: Gen. Relativ. Gravit. **42**, 2643 (2010a)
Sharif, M., Shamir, M.F.: Mod. Phys. Lett. A **25**, 128 (2010b)
Sharif, M., Zubair, M.: Astrophys. Space Sci. **330**, 399 (2010a)
Sharif, M., Zubair, M.: Int. J. Mod. Phys. D **19**, 1957 (2010b)
Sharif, M., Zubair, M.: Astrophys. Space Sci. **339**, 45 (2012a)
Sharif, M., Zubair, M.: Astrophys. Space Sci. **342**, 511 (2012b)
Singh, C.P.: Pramāna **72**, 429 (2009a)
Singh, C.P.: Gravit. Cosmol. **15**, 381 (2009b)
Singh, C.P., Beesham, A.: Int. J. Mod. Phys. A **25**, 3825 (2010)
Singh, C.P., Kumar, S.: Int. J. Mod. Phys. D **15**, 419 (2006)
Singh, C.P., Zeyauddin, M., Ram, S.: Int. J. Theor. Phys. **47**, 3162 (2008)
Spergel, D.N., et al.: Astrophys. J. Suppl. Ser. **148**, 175 (2003)
Steinhardt, P.J., Wang, L.M., Zlatev, I.: Phys. Rev. D **59**, 123504 (1999)
Tegmark, M., et al.: Phys. Rev. D **69**, 103501 (2004)
Tonry, J.L., et al.: Astrophys. J. **594**, 1 (2003)
Yilmaz, İ., Baysal, H., Aktaş, C.: Gen. Relativ. Gravit. **44**, 2313 (2012)
Zlatev, I., Wang, L.M., Steinhardt, P.J.: Phys. Rev. Lett. **82**, 896 (1999)

Geotechnical Behaviour of Fine Sand Mixed with Pond Ash and Lime

Shilpi Gupta[1], Prateek Negi[2], Ashok Gupta[3]

Abstract: The present work has been carried out to study the effect of lime and pond ash on geotechnical behaviour of fine sand. Fine sand found in deserts and near sea shores is not readily suitable for supporting buildings and roads, as it is loose and vulnerable to collapse upon wetting. Fine sand having high permeability and low shear strength is also not suitable for earthen dam construction. For the same reasons, fine sand is not suitable for constructing landfill liners and covers. The disposal of pond ash from power plants has been a major environmental concern over a period of last few decades. In presence of moisture, pond ash reacts with lime and forms cementitious compounds which can be used to increase the strength of weak soil.

A series of testing including scanning electron micrographs and x-ray diffraction has been performed to characterize the materials. Various proportions of mixes of fine sand with pond ash keeping lime constant were tested to find out the optimum percentage of pond ash.

Keywords: Fine Sand, Lime, Pond Ash, Scanning electron micrographs, X-ray diffractograms, Soil compaction, Shear strength, Triaxial test.

1 INTRODUCTION

SAND is a natural granular material composed of finely divided rocks and mineral. The composition of sand varies according to the parent rock and the mode of transportation. Sand is transported by various agents like wind and water, and is deposited in the form of beaches, dunes, sand spits, sand bars etc.

On the basis of particle size, sand can be classified as following:

- Fine sand = 0.075 to 0.425 mm
- Medium sand = 0.425 to 2 mm
- Coarse Sand = 2.0 to 4.75 mm

Fine sand obtained from sea shores is generally brown in color and angular in shape. It contains salts with it which attracts moisture from atmosphere. Such absorption causes dampness and disintegration of work in which it is used. Sea sand also retards setting action of cement. Due to these reasons, this kind of fine sand is generally avoided for engineering purposes. It is used only as a local material for nonstructural purposes.

Desert sand is also usually fine-grained and poorly graded sand, which create major problems for construction of infrastructures. Fine sand is also not suitable for dam construction because of its relatively high permeability and low shear strength. For the above reasons, fine sand is not suitable for construction of landfill liners and covers.

Pond ash is a waste product from thermal power industry. More than 150 million tones of pond ash is produced annually in India. The un-utilized fly ash and bottom ash are taken to ash ponds for deposition. The ash deposited in the ash pond is known as pond ash.

Today around 20,000 hectares of land is occupied by ash ponds. It not used in some manner, pond ash will become an increasing burden on the environment. Now days it is not regarded as waste material.

In presence of moisture, pond ash chemically reacts with lime at ordinary temperature and forms a compound possessing cementitious properties. After reactions between lime and pond ash, calcium silicate hydrates are formed which are responsible for the high strength of the new formed compound. **S. Bhuvaneshwari et al. [1].**

2 LITERATURE REVIEW

E.A. Mohamedzein Yahia et al. (2006) has carried out a study on stabilization of desert sands with incinerator ash for geotechnical engineering applications and concluded that there is a significant increase in unconfined compressive strength, and shear strength parameters and a reduction in permeability with increase in ash *percentage. The bearing capacity provided by such* mixtures will be adequate enough to support a low-to-moderate high rise

1 Post Graduate Student (Geotechnical Engineering,Civil Engg. Dept, DTU, New Delhi, India),PH- 08802483044. E-mail:gupta.shilpi33@gmail.com

2 Asst. Prof. (Civil Engg. Dept, Galgotias University, Greater Noida, Uttar Pradesh, India)· PH- 09953919464. E-mail: negidynamic@gmail.com

3 Professor & Head (Civil Engg. Dept, Delhi Technological University, New Delhi, India), PH-09313505567. E-mail: akumargupta@dce.ac.in

building. The mixture can also be used in sub grade construction. The hydraulic conductivity of the mixture is low and the mixture can be used as a landfill cover.

Erdal Cokca (2001) studied the effect of fly ash on expansive soil and his experimental findings confirmed that the plasticity index, activity and swelling potential of the samples decrease with increasing percentage of stabilizer. The undrained shear strength of expansive soil with fly ash increases with increase in fly ash.
A similar study was carried out by **Phani kumar and Sharma (2004)** and **S. Bhuvaneshwari et al. (2005).**

N.S. Pandian (2004). The detailed investigations carried out on fly ash elsewhere as well as at the Indian Institute of Science showed that fly ash has good potential for use in geotechnical applications. Its low specific gravity, freely draining nature, ease of compaction, insensitiveness to changes in moisture content, good frictional properties, etc. can be gainfully exploited in the construction of embankments, roads, reclamation of low-lying areas, fill behind retaining structures, etc. It can be also used in reinforced concrete construction since the alkaline nature will not corrode steel. This not only solves the problems associated with the disposal of fly ash (like requirement of precious land, environmental pollution, etc.) but also helps in conserving the precious top soil required for growing food.

3 Experimental Works

3.1 Materials used

A sample of fine sand was collected from beech in Chennai region, Delhi and sample of pond ash was collected from Rajghat power station ash pond, Delhi. Properties of fine sand and pond ash are shown in Table 1 below.

TABLE 1
Properties of Fine Sand & Pond Ash

Properties	Fine sand	Pond ash
Specific gravity	2.65	2.11
Plastic Limit	Non-Plastic	Non-Plastic
Optimum moisture content (%)	8.77	18.20
Maximum dry density (g/cc)	1.58	1.21
Un-soaked CBR (%)	7.17	11.95
Cohesion (kPa)	14.34	0.0
Angle of internal friction (deg.)	36.02	28.2

3.2 Admixture proportions and tests conducted

Lime was added as 3%, 6%, 9% and 12% by weight to fine sand to find out the optimum value of lime. Later on keeping the amount of lime fixed at 9%, pond ash was added in proportions of 8%, 12%, 16% and 20% by weight. The following experiments were performed on the fine sand mixed at different mixes of fine sand, lime and pond ash.

The formation of bonds was observed under scanning electron microscope (SEM) and mineral composition of materials was found with x-ray diffraction technique (XRD). Specific gravity test was conducted as per IS: 2720 (Part 3)-1980. Particle size distribution was conducted as per IS: 2720 (Part 4) -1985. Standard Proctor Test was conducted as per IS: 2720 (Part 7) -1980. California Bearing Ratio (CBR) was conducted as per IS: 2720 (Part 16) -1987. UU triaxial test was conducted as per IS: 2720 (Part 11) -1983.

3.3 Results and Discussion

Scanning Electron Microscope tells us about the morphology of the sample. SEM micrographs of lime-fine sand mixtures and pond ash-lime-fine sand mixtures are shown from fig. 14 to 24. SEM micrographs showed better interlocking between fine sand, pond ash particles and lime. Hence mixing of pond ash and lime in fine sand enhances the strength of the mixture.

3.4 X- Ray Diffraction

X-ray diffraction technique reveals information about the crystal structure, chemical composition, and physical properties of materials. The x-ray diffraction of dry sample of fine sand, pond ash and lime was conducted and X-ray diffractograms are shown from fig. 25 to 27. The XRD result shows the presence of quartz, mullite in fine sand. The peaks in XRD of the pond ash showed the presence: silica (SiO_2), mullite ($A1_6Si_20_{13}$), magnetite (Fe_3O_4) and presence of calcite in lime, when matched with JCPDS (*Joint Committee on Powder Diffraction Standards*) data book.

3.5 Particle Size Distribution

The result of particle size distribution is shown in fig. 1. The result showed that fine sand is uniformly graded where as pond ash is poorly graded. Their particle size distribution curve is very much parallel which shows their suitability to be mixed well.

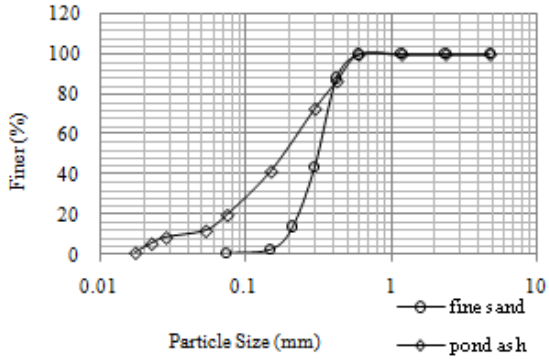


Fig. 1: Comparison of particle size analysis

3.6 Light compaction test

The variation in maximum dry density and optimum moisture content for lime-fine sand mixture is represented in fig. 2 and 3 and variation for pond ash-lime-fine sand mixture is represented in fig. 4 and 5. MDD and OMC are increasing with increase in lime percentage in fine sand. MDD is decreasing with increase in pond ash percentage in fine sand but it is greater than MDD obtained for fine sand with 9% lime. OMC first decreases when pond ash is added and on further addition of pond ash, OMC starts increasing.

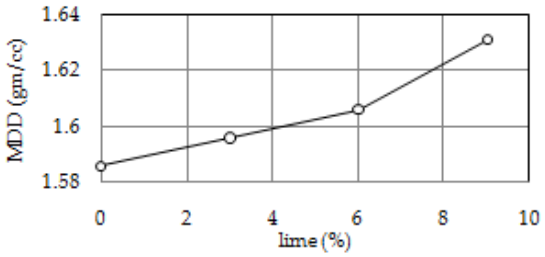


Fig. 2: Influence of lime on MDD

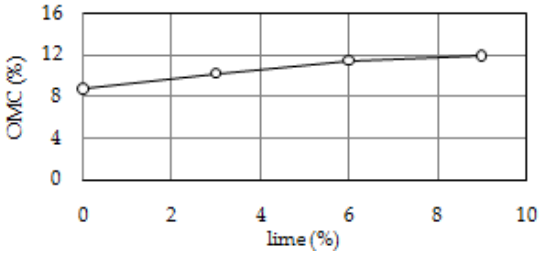


Fig. 3: Influence of lime on OMC

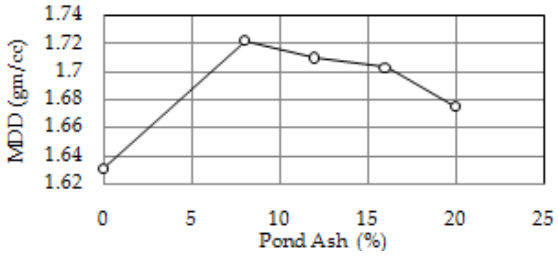


Fig. 4: Influence of pond ash and lime on MDD

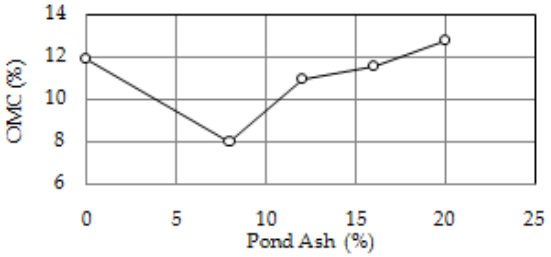


Fig. 5: Influence of pond ash and lime on OMC

3.7 Unsoaked CBR Test

The CBR test results are showed in fig. 6 to 9.CBR Test results reveals that CBR value increases every time when lime percentage is increased and CBR value increases every time when pond ash percentage is increased. But it is less than CBR value obtained for fine sand with 9% lime.

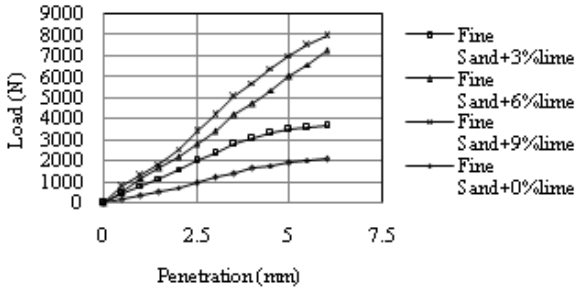


Fig. 6: Influence of lime on CBR curves

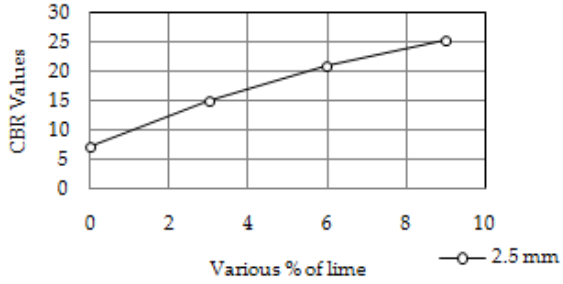


Fig. 7: Influence of lime on CBR values

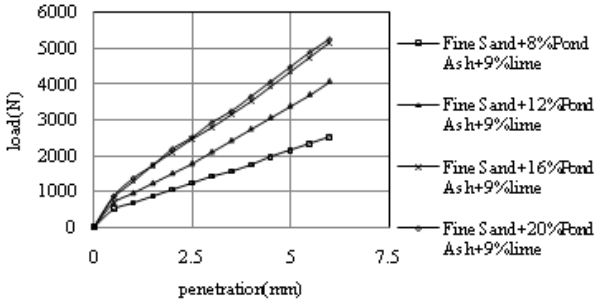


Fig. 8: Influence of pond ash and lime on CBR curves

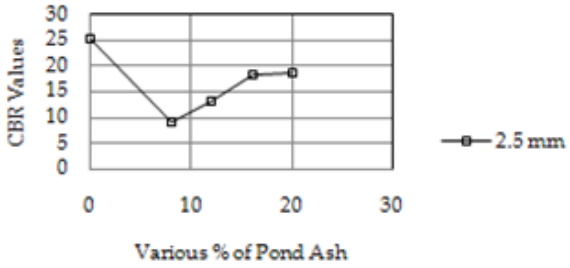


Fig. 9: Influence of pond ash and lime on CBR values

3.8 UU Triaxial test

The UU triaxial test results are represented in fig. 10 to 13.UU test results reveals that cohesion(c) and angle of internal friction(ϕ) increases with increase in lime percentage in fine sand. Up to 9% lime after that it is decreases. And with addition of pond ash with 9% lime cohesion and angle of internal friction further increases up to 16% pond ash.

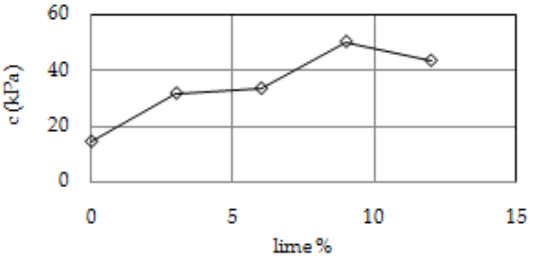


Fig. 10: Influence of lime on cohesion

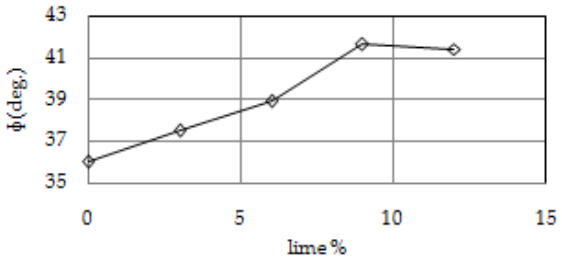


Fig. 11: Influence of lime on angle of internal friction

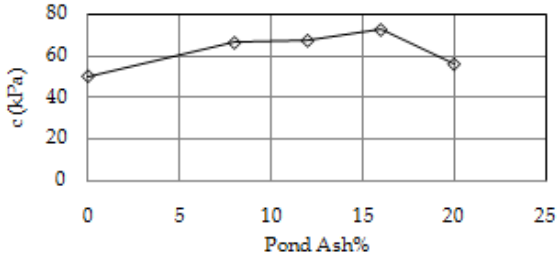


Fig. 12: Influence of pond ash and lime on cohesion

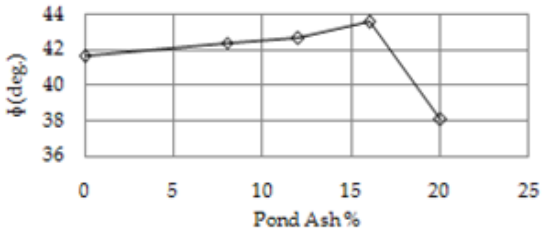


Fig. 13: Influence of pond ash and lime on angle of internal friction

4 Conclusions

1. Scanning electron micrographs tells us about the morphology of the sample mixtures. Fine Sand particles can be seen with sharp edges. The SEM of pond ash shows the spherules of alumina silicates. Dark matter present shows magnetite. SEM results shows better interlocking between fine sand, pond ash particles and lime. Hence mixing of pond ash and lime in fine sand can result in enhanced strength of fine sand.
2. The X-Ray Diffraction results shows the presence of quartz, mullite in fine sand, the presence of silica, mullite, magnetite in pond ash and presence of calcite in lime.
3. The density bottle tests reveal that the specific gravity of the fine sand is 2.65 and of pond ash is 2.11
4. The results from the grain size analysis reveal that fine sand is uniformly graded and pond ash is not well graded, though in particle size distribution they were very much parallel.
5. The standard proctor test indicates that MDD of fine sand increases with lime percentage. It can be due to interlocking of lime particles in voids of fine sand. Also, MDD of fine sand increases with pond ash percentage up to 8% after that it decreases. The reason may be , at first pond ash particles fill the voids present in fine sand but after that pond ash

particles(which are lighter) will replace the fine sand particles(which are heavier than pond ash particles) and hence, reduces the MDD.

6. The CBR test reveals that CBR value increases every time when lime percentage is increased. This can be due to the interlocking of lime particles in fine sand. Also, lime provides strength on hydration and acts as a binder between particles. CBR value increases with increases in pond ash percentage due to cementitious material formed by pond ash and lime.
7. The triaxial test reveals that cohesion and angle of internal friction of fine sand increases with lime % up to 9% after that they decreases. This can be because of increase in slippage between fine sand particles due to extra lime. Addition of pond ash with 9% lime in fine ash shows the cohesion and internal friction increases with pond ash percentage up to 16% pond ash. Cohesion and angle of internal friction of fine sand maximizes at 16% pond ash and 9% lime.

ACKNOWLEDGMENT

I would like to express my deepest sense of gratitude and indebtedness to God and my parents for their encouragement and persistent support which has helped me to do better in all of my endeavors.

I would also like to thank my guide and motivator Prof. A.K. Gupta, H.O.D, Civil Engineering Department, Delhi Technological University for his valuable guidance and support in all the phases from conceptualization to experimental and final completion of the project.

I wish to convey my sincere gratitude to my senior sir Prateek Negi who has enlightened me during my project.

I would also like to thank all the faculties of Civil Engineering Department, Delhi Technological University.

I am deeply thankful towards all the lab assistants of my college who have helped me conduct the experiments.
I thank all the people directly or indirectly involved in successful completion of this project.

REFERENCES

1. Bhuvaneshwari, S., Robinson R. G. and Gandhi, S. R.: "Stabilization o f Expansive Soils Using Fly Ash". Fly Ash India, New Delhi, India, 2005.
2. Cokca, E.: ''Use of class C fly ash for the stabilization of an expansive soil''. J. Geotech. Geoenviron. Eng., 127(7), 568–573, 2001.
3. Diouf, B., Skidmore, E.L., Layton, J.B., Hagen, L.J. : "Stabilizing fine sand by adding clay: laboratory wind tunnel study". Soil technology, vol-3, pp. 21-31, 1990.
4. Fei Liu, Junshu Wu, Kunfeng Chen and Dongfeng Xue, "Morphology Study by Using Scanning Electron Microscopy". Microscopy: Science, Technology, Applications and Education.
5. IS: 2720 (part 3): IS code for determination of specific gravity, 1980.
6. IS: 2720 (part 4): IS code for determination of particle size, 1985.
7. IS: 2720 (part 5): IS code for determination of liquid limit and plastic limit– 1983, 1985.
8. IS: 2720 (part 7): IS code for determination of water content – dry density relation using light compaction, 1983.

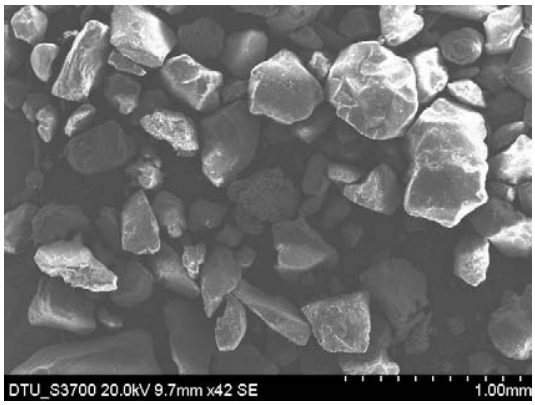


Fig. 14: Fine Sand at 1mm scale

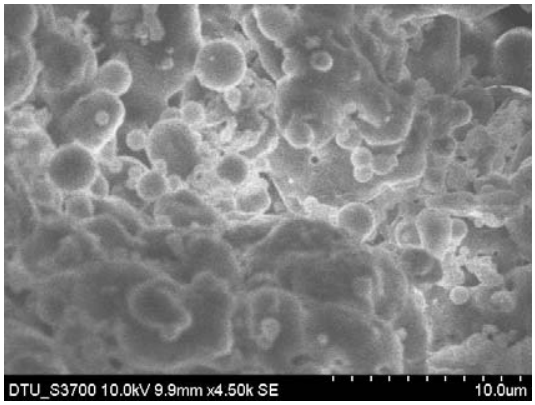


Fig. 15: Pond Ash at 10 µm scale

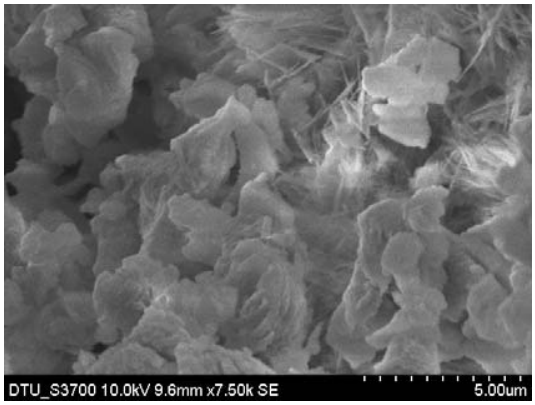


Fig. 16: Lime at 5 µm scale.

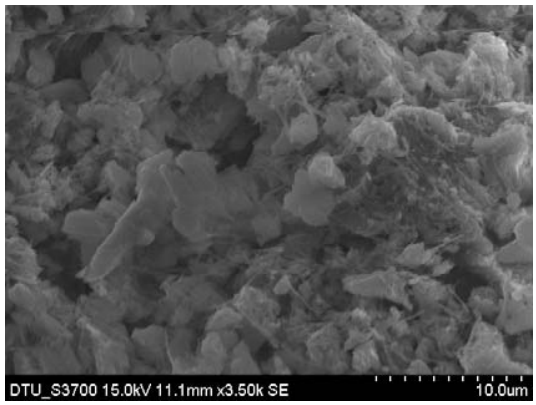


Fig. 17: Fine Sand+ 3% Lime at 10 µm scale

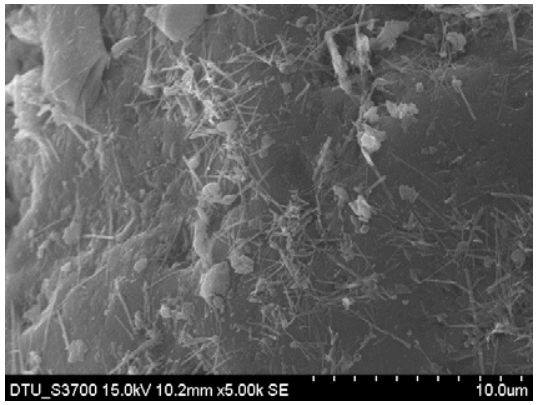


Fig. 18: Fine Sand+ 6% Lime at 10 µm scale

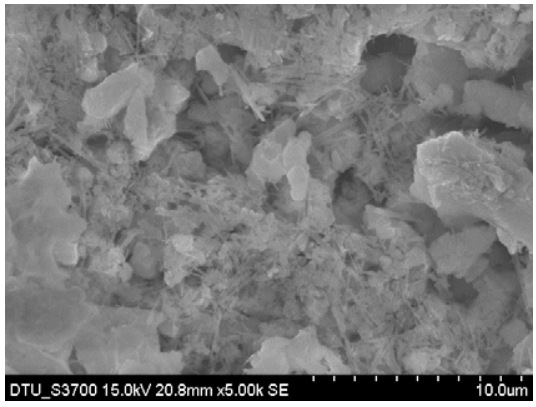


Fig. 19: Fine Sand+ 9% Lime at 10 µm scale

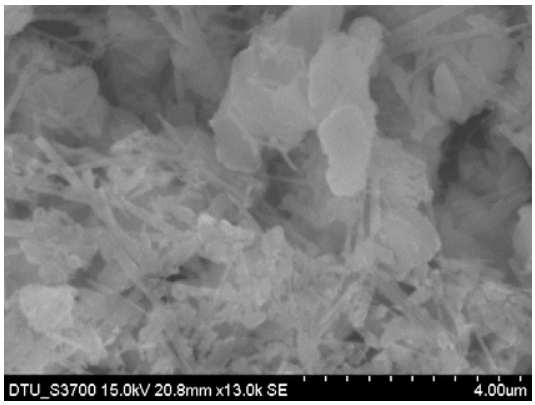


Fig. 20: Fine Sand+ 12% Lime at 4 μm scale

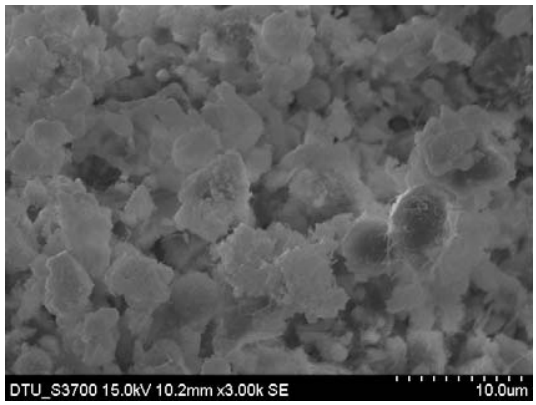


Fig.23: Fine Sand+16% Pond Ash +9% Lime at 10 μm scale

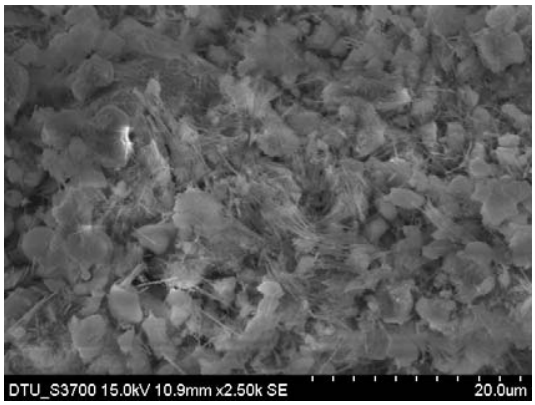


Fig. 21: Fine Sand+8% Pond Ash +9% Lime at 20 μm scale

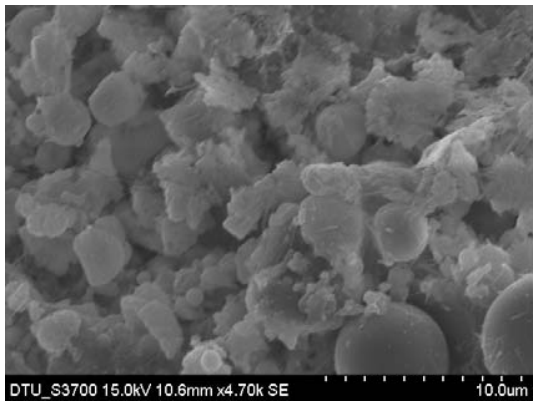


Fig. 24: Fine Sand+20% Pond Ash +9% Lime at 10 μm scale

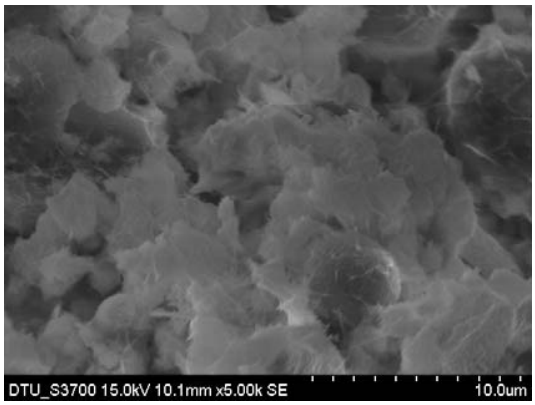


Fig. 22: Fine Sand+12% Pond Ash +9% Lime at 10 μm scale.

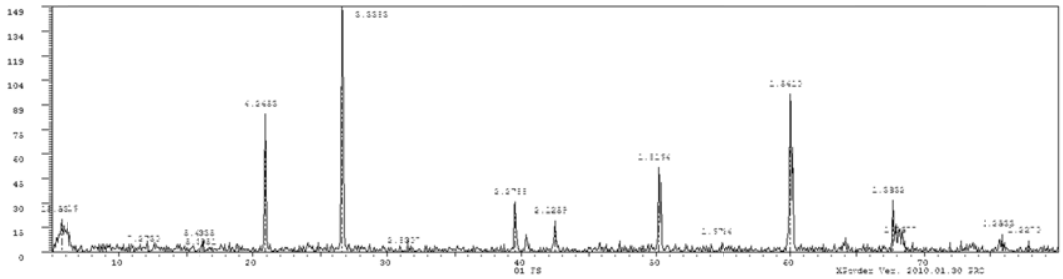


Fig.25 XRD pattern of fine sand particles

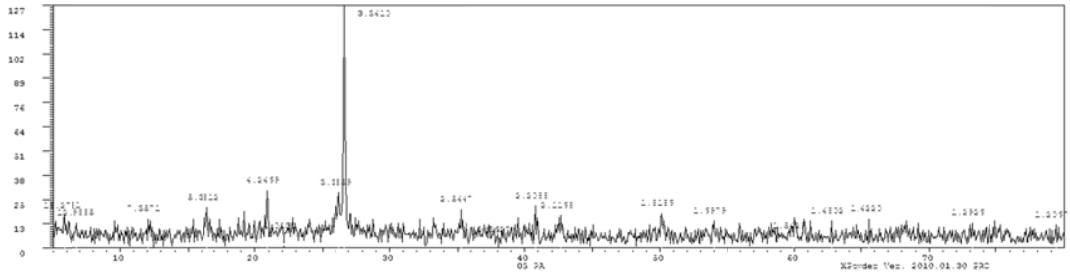


Fig.26 XRD pattern of pond ash particles

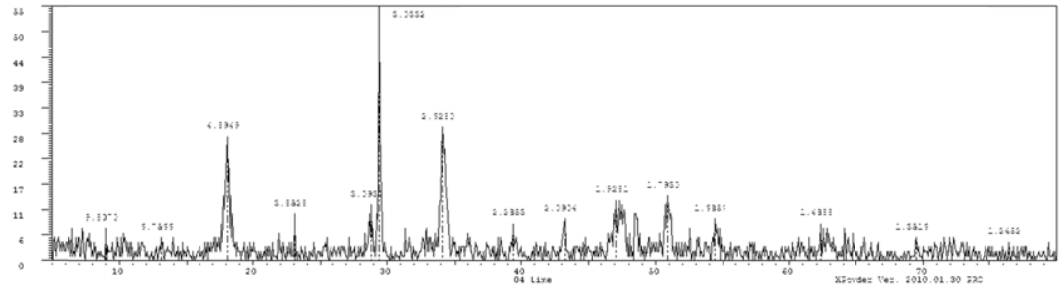


Fig.27 XRD pattern of lime

High Performance Concrete

Abhishek Anand, Abhinav Daalia

Civil Engineering, Civil Engineering, Delhi Technological University, Delhi Technological University
Main Bawana Road, New Delhi, India
abhi21anand@gmail.com

abhinav.daalia@gmail.com

Abstract*— The current trend of use of superlatives in concrete technology may strike as somewhat disconcerting to many .We had high strength concrete, hyperplasticiser, and superplasticisers, very reactive Pozzolana, and now high performance concrete. It is difficult to imagine any concrete being manufactured and used, which is not intended to perform to the extent; high performance concrete is not a new material of construction. It is difficult to imagine any concrete being manufacture and used, which is not intended to perform. The only difference is the level of performance, which is higher than ordinary. High Performance Lightweight Concrete (HPLC) has been extensively investigated for, among other applications, use in oil drilling platforms in severe environments.The relationships and allowable stresses and the stress block given in structural codes for normal strength concrete (e.g. IS 456 or IRC 21) will require modification. Acceptance testing on site has to be more than cube testing at28 days. Where durability of concrete is the driving force for adoption of high performance concrete, in-situ permeability tests are performed as a matter of routine.

***Keywords*— High Performance Concrete, Durability, Superplasticisers, Silica Fume, Fly-ash, Permeability, Workability, Aggregate**

I. INTRODUCTION

The current trend of use of superlatives in concrete technology may strike as somewhat disconcerting to many .We had high strength concrete, superplasticisers and hyper plasticiser, very reactive pozzolana, and now high performance concrete. It is difficult to imagine any concrete being manufacture and used, which is not intended to perform. To that extend, high performance concrete is not a new type of material of construction. The only difference is of performance, which is higher than ordinary. . High Performance Lightweight Concrete (HPLC) has been extensively investigated for, among other applications, use in oil drilling platforms in severe environments.

II. DEFINITIONS

Many attempts have been made to define high performance concrete .A quantitative definition is that

- It should have a maximum water cement ratio of .35.a minimum durability factor of 80% in freeze-thaw resistance test as per ASTM C666,a minimum compressive strength of 21Mpa at 4hours ,34Mpa at 1day ,or 69Mpa at 28days.

However, such quantitative definitions may not satisfactory in all situations. Among general qualitative definitions is

- High performance concrete is defined as concrete which meets special performance and uniformity requirements that cannot always be achieved routinely by using only conventional materials and normal mixing, placing, and curing practices. The requirements may involve enhancements of characteristics such as placement and compaction without segregation, long-term mechanical properties, early-age strength, toughness, volume stability, or service life in severe environments.

For high strength, water cement ratio should below. The strength water-cement ratio rule holds good for concrete strength of about 100Mpa or more. Low water-cement ratio is also required for low permeability of concrete, which is vital for high durability

III. CHOICE OF MATERIALS

The ingredients of high performance concrete are to be chosen in the desired performance and yet ensure economy. In this process, the materials are stretched to their limits of performance.

A. CEMENT

It should have high strength, often with early strength OPC53 grade is chosen, although blended cements like Portland pozzolana cement or Portland slag cement, or even ternary blends are preferred in others countries, for high durability. In India, 70% slag or 30% processed fly ash mixed with OPC at the site mixer has been used. Low C3A in OPC is preferable.

B. AGGREGATE

Selection of aggregate was done carefully after conducting several tests in the laboratory on samples obtain from various sources, aggregate should be strong, clean and equi dimensional, i.e not flaky or elongated . For strength grades above M60, nominal maximum size of 20 mm is preferable. Natural sand of zone II or III of IS 383, having fineness modulus of 2.6 or more, is suitable.

C. FLY-ASH

Should conform to IS;3812-Part-I-2003, and should be finer than cement .Uniformity is very important. Source of fly ash

should be chosen carefully. Processed fly-ash, in which, the fine fraction is developed by controlled grinding, is preferable.

D. *SUPERPLASTICISERS*

Selection of appropriate quality of superplasticiser is crucial in high performance concrete. Trails were conducted in the laboratory on samples obtained from various superplasticisers with cement was also studied. With some of the superplasticisers tried out the slump retention was satisfactory but there was a reduction in the one –day compressive strength to the extent of around 25% from the desired target value. Finally, a tailor-made superplasticiser was used which gave desired slump retention without the loss of strength. The superplasticiser should conform to IS-9103.Some properties of superplasticisers are:

DESCRIPTION	**VALUE**
PH	8.2
CHLORIDE PERCENT	.016
SULPAHTE PERCENT	.024
SOLID CONTENT PERCENT	38.5
SPECEFIC GRAVITY	1.224

Some common types are:

1 .Modified Lignosulphonates (MLS).
2. Melamine Sulphonates (MSF).
3. Napthalene Sulphonates (NSF).
4. Acrylic Polymer (AP) based.
5. Cross linked Acrylic Polymer (CLAP).
6. Polyacrylates

Combinations of the above are also common .In general efficient superplasticiser allow water reduction of at least 20% and 30% or more with new generation products, enhance slump retention and also reduce slump loss with time. Compatibility with the cement is very necessary.

E. GROUND SLAGS

There is no BIS specification for ground slag as additive to concrete. Provisions of IS-12089 (granulated slag for manufacture of slag cement) are adequate.

F. SILICA-FUME

A by product resulting from the reduction of high purity quartz with coal in electric furnaces at temperature of 2000ºC in the production of Ferro-silicon alloy or silicon metal. In the process of ,SiO-vapours are produced ,which oxidize and condense in the form of very tiny spheres of non –crystalline silica. The use of silica fume was specified mainly for ensuring the long term durability of structure IS-155388-2003.The properties of silica fume are

S.No	DESCRIPTION	VALUE
1.	SiO2	9.1
2.	H20	0.5
3.	Loss on percent, ignition	2.4
4.	Carbon percent	1.8
5.	Density, kg/m3	575

G. WATER

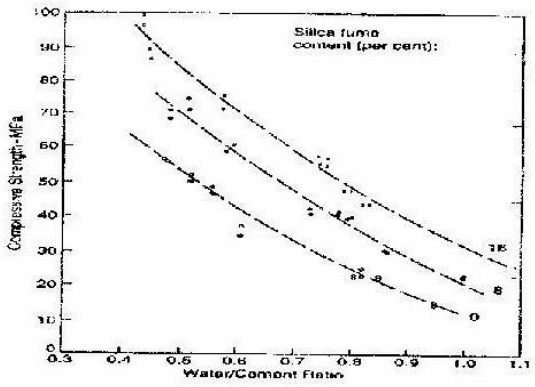
In high performance concrete the water used for mixing and curing was also obtained from a nearby local source the and requirements given in IS-456:2000.Properties of the water sample are given:

S.No.	DESCRIPTION	VALUE
1.	Source	Local Well
2.	Organic Solids, gm/ml	88
3.	Inorganic Solids, gm/ml	242
4.	Sulphates as SO4, gm/ml	21
5.	Chlorides as Cl, gm/ml	19
6.	Suspended Matter, gm/ml	1.8
7.	PH	7.6
8.	Alkalinity	6.2 ml required to neutralize 200-ml water using 0.1 N HCl

IV. MECHANISM OF HIGH-PERFORMANCE

High strength and low permeability are logical development of presence of silica fume and superplasticisers in concrete the dual requirements of high strength and low permeability are linked to each other through the need to reduce the volume of relatively larger capillary pores. As pointed out earlier, this is achieved by low water-cement ratio as well as dense packing of fine particles.The role of superplasticisers, long chain molecule organic compound, is to get adsorbed on to cement grain, impart a negative charge to them, which repel each other and get deflocculated and disperse. the resulting improvement in workability of concrete could be either to flowing concrete for same water& cement contents as in the control mix.

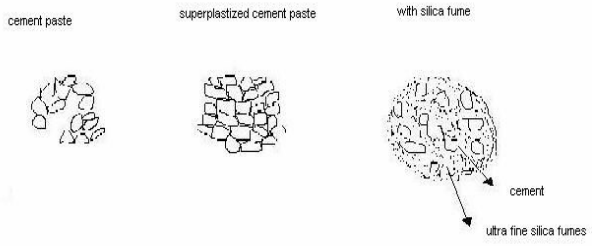
Alternatively, it enables water content to be reduced by 20% or more and results in high strength, because of low water-cement ratio.



The role of silica is many fold because of enormous surface area as well as relatively large content of glassy silica, it is very reactive pozzolana. When contribution of silica fume to compressive strength of concrete is compared based on water - cement ratio in mix, mixes containing silica has high compressive strength at 28days(fig). The effect of silica fume is greater than that of cement replaced; the 'efficiency factor' is about three. In other words, 20Kg of silica fumes can be replace about 60Kg of cement and yield same strength. It also helps in augmenting early strength of concrete. Since the heats of hydration of both are of same order, incorporation of silica fume enables the heat rises in concrete to be lowered -a critical advantage for mass concrete. However, incorporation of silica fume in concrete increases the water demand. Hence, superplasticisers are required.

Dense packing is another basis of superior performance of concrete containing silica fume and superplasticiser. The cement grains, which tend to flock together, are dispersed by superplasticiser. The extremely fine silica fume particles are then packed in the space between dispersed cement grains and normally packed fine and coarse aggregate. The mechanism is schematically shown in (fig). The overall result is a denser microstructure. The concrete exhibit less porosity with no evidence of capillary pores and only very narrow less than 0.5mm gel pores is visible under high magnification in electron microscope. C-S-H gel particles in concrete containing silica fumes appear not as individual particles, but rather as a massive, dense structure. By residing in the pores in hydrated cement paste, silica fumes particles, on hydration, block the pores. Such pore -refining action reduces the size of pores,although the overall porosity may remain the same shows some results

Fig. Packing of cement paste containing Superplasticiser and silica fume



Another important mechanism is improving the transition zone around aggregate particles. In normal strength concrete with only cement, the transition zone around aggregate is 20mm to 100mm wide and richer in calcium hydroxide and ettringite, as against C-S-H phase in the bulk matrix. The porosity is also higher. Thus, the transition zone forms a weak link. In presence of silica fume, dense C-S-H occupies all the space around aggregate and direct bond with aggregate is established. Result of strengthened transition zone is reduction in microcracking at the interface between cement paste and aggregate. The stress-strain curve remains linear up to about 85 % of failure stress or higher.

V. TYPES OF HIGH PERFORMANCE CONCRETE

A. HIGH WORKABILITY CONCRETE
B. LIGHT WEIGHT CONCRETE
C. HIGH STRENGTH CONCRETE
D. SPRAYED CONCRETE
E. ROLLER CONCRETE

A. HIGH WORKABILITY CONCRETE

The workability of fresh concrete should be suitable for each specific application to ensure that the operations of handling, placing and compaction can be undertaken efficiently. The handling and placing properties of concrete mixes can be improved considerably by the use of cement replacement materials such as pulverized fuel ash or ground granulated blastfurnace slag. Further more, the use of admixtures such as water reducers and superplasticisers have beneficial effects on workability without compromising other concrete properties.

On site productivity can be greatly increased by utilizing highly workable concretes. They are especially suitable in the following applications:

- Inaccessible locations
- Large flat areas
- Underwater applications
- Pumping concrete over long distances

The high workability concrete is further divided into the following types

1. Self-compacting concrete
2. Pumped concrete
3. Foamed Concrete

1. Self- Compacting Concrete

SCC has been described as the most revolutionary development in concrete construction for several decades. Originally developed to offset a growing shortage of skilled labor, it has proved beneficial economically because of a number of factors, including:

- faster construction
- reduction in site manpower
- better surface finishes
- easier placing
- improved durability

- greater freedom in design
- thinner concrete sections
- reduced noise levels,absence of vibration
- safer working environment

2. Pumping

Pumping is one of the most popular techniques worldwide to transport fresh concrete. When placing large quantities of fresh concrete, piston pumps are generally used. Concrete is pushed alternately by two pistons acting in cylinders.

3. Foamed Concrete

Foamed concrete is a highly workable, low-density material which can incorporate up to 50% entrained air. It is generally self-leveling, self-compacting and may be pumped. Foamed concrete is ideal for filling redundant voids such as disused fuel tanks, sewer systems, pipelines, and culverts - particularly where access is difficult. It is a recognized medium for the reinstatement of temporary road trenches. Good thermal insulation properties make foamed concrete also suitable for sub-screeds and filling under-floor voids.

B. LIGHT WEIGHT CONCRETE

Lightweight concretes can either be Lightweight Aggregate concrete, Foamed concrete or Autoclaved Aerated concrete (AAC). Such lightweight concrete blocks are often used in house construction.

1. Lightweight Aggregate Concrete:

Lightweight aggregate concrete can be produced using a variety of lightweight aggregates.Lightweight aggregates originate from either:

- Natural materials like volcanic pumice
- The thermal treatment of natural raw materials like clay, slate or shale i.e. Leca
- Manufacture from industrial by-products such as fly ash, i.e. Lytag
- Processing of industrial by-products like FBA or slag

The required properties of the lightweight concrete will have a bearing on the best type of lightweight aggregate to use.If little structural requirement, but high thermal insulation properties are needed, then a light,weak aggregate can be used. This will result in relatively low strength concrete. Lightweight aggregate concretes can however be used for structural applications, with strengths equivalent to normal weight concrete.

The benefits of using lightweight aggregate concrete include:

• Reduction in dead loads making savings in foundations and reinforcements.
• Improved thermal properties.
• Improved fire resistance.
• Savings in transporting and handling precast units on site.
• Reduction in formwork and propping

2. Foamed Concrete

Foamed concrete is a highly workable, low-density material which can incorporate up to 50% entrained air. It is generally self-leveling, self-compacting and may be pumped. Foamed concrete is ideal for filling redundant voids such as disused fuel tanks, sewer systems, pipelines, and culverts - particularly where access is difficult. It is a recognized medium for the reinstatement of temporary road trenches. Good thermal insulation properties make foamed concrete also suitable for sub-screeds and filling under-floor voids.

3. Autoclaved- Aerated concrete (AAC)

AAC was first commercially produced in 1923 in Sweden. Since then, AAC construction systems such as masonry units, reinforced floor/roof and wall panels and lintels have been used on all continents and every climatic condition. AAC can also be sawn by hand, sculpted and penetrated by nails, screws and fixings.

C. HIGH STRENGTH CONCRETE

The definition of high strength concretes is continually developing. In the 1950s 34N was considered high strength, and in the 1960s compressive strengths of up to 52N were being used commercially. More recently, compressive strengths approaching 138N have been used in cast-in-place buildings.High-strength concrete columns can hold more weight and therefore be made slimmer than regular strength concrete columns, which allows for more useable space, especially in the lower floors of buildings.

D. SPRAYED CONCRETE

There are many applications where traditional formwork is not a suitable method of supporting concrete and so sprayed concretes can be used.

These applications include:

• Shell roofs and domes
• Retaining walls
• Piled wall facings
• Silo structures
• Diaphragm walls
• Blast proof structures
• Bank vaults
• Underground Construction
• Tunnel linings
• Free form structures such as those seen in theme parks
• Water Retaining Structures

• Sea and river walls

• Reservoirs and dams

• Aqueducts Swimming pools

E. ROLLER COMPACTED CONCRETE

Roller compacted concrete, a durable paving material that carries heavy loads. It is now developing as a fast, economical construction method for dams, off-highway pavement projects, heavy-duty parking and storage areas, and as a base for conventional pavement. It is a stiff, no-slump concrete mixture with the consistency of damp gravel comprised of local aggregates or crushed recycled concrete, portland cement, and water. The mixture is placed and roller compacted with the same commonly available equipment used for asphalt pavement construction. The process requires no forms, finishing, surface texturing, or joint sawing and sealing.

VI. EXAMPLES OF APPLICATION IN INDIA

The applications of high performance concrete in are listed below

A. Infrastructure Projects

- Bandra-Worli Sea link:

- M60 grade performance concrete for superstructure was specified.
-Target 28-days compressive srength-74N/mm.
-Slump obtained was 120mm,28-day compressive strength in the field was 75.16N/mm².

Mix Design for Bandra Worli Sea-Link

Cement	330 Kgs
Silica Fume	33 Kgs
Coarse Aggregate	1028 Kgs
Fine Aggregate	1026 Kgs
Water	130.68 (w/c=0.34, w/b=0.31)
Superplasticiser	10 Kgs

(*All quantities are per m³ of concrete)

- JJ Flyover, Mumbai

-M7O grade was supplied from a readymade mix concrete plant.
-Target strength-83.2Mpa
-Slump required was 130 to 180 mm at the RMC plant, and 80 to 120mm at the site of placement after 150mm.
-Field strength obtained were –79.6Mpa at 28 days and 94 Mpa at 365days.
-Chloride ion permeability(ASTM 1202)-100 Coulombs (very low)

Mix design for JJ Flyover, Mumbai

Cement	500 Kgs
Silica Fume	50 Kgs
Coarse Aggregate	682 Kgs
Fine Aggregate	1156 Kgs
Water	148 litres (w/b=0.31)
Superplasticiser	10 Kgs

(*All quantities in per m³ of concrete)

B. Nuclear Power Projects

- Kaiga Atomic Power Project (Kaiga-unit-1 and 2)
- Rajasthan Atomic Power Plant (RAPP-unit-3and 4)
- Tarapur Atomic Power Project (TAPP-unit 3 and 4)

High performances concrete for these applications were required to have moderate compressive and high tensile strength; very high durability; low creep and shrinkage; low permeability and good workability.

C. Water Resource Sectors

High-performance concrete has been used in hydraulic structures, essentially to redder high abrasion-erosion resistance. In this process, the replacements of considerable cement quantity by silica fume. The two major areas of application are:

- High performance concrete with high strength and wear resistance
- Having improved bond to other materials, is achieved with incorporation of silica fume and low water/cement ratio.

Materials,Kg (Per m3 concrete)	Kaiga (IC Dome)	RAPP(IC DOME)	TAPP(IC STRUCTURE)
Cement	475	475	475
Silica Fume	35.6	36	36.6
Water/Ice	163	152	152
Coarse Aggregate	1092	1047	1133
Fine Aggregate	695	730	721
Superplasticisers,ltrs	8.4	9.63	9.53
Water/Cement	0.343	0.32	0.32
Water/Binder	0.32	0.3	0.3

Superplasticisers are required to make such concrete workable. Application in India include:

- M70 grade concrete for spillways, stilling basin in Tehri Dam
- M50 grade concrete for spillways, stilling basin and flushing tunnels in Baglihar Hydroelectric

Project.Chamara- II and Tala Hydroelectric project (Butan).

- As well as for repair of stilling basins in Sri Sailam Project and Salal Dam.

VII. CONCLUSIONS

Because of altered microstructure and improved transition zone, engineering properties like tensile strength, modulus of rupture , elastic modulus etc. as functions of compressive strength as well as limi6ting ultimate strain are different in case of high performance concrete. The relationships and allowable stresses and the stress block given in structural codes for normal strength concrete (e.g. IS 456 or IRC 21) will require modification. Acceptance testing on site has to be more than cube testing at28 days. Where durability of concrete is the driving force for adoption of high performance concrete, in-situ permeability tests are performed as a matter of routine. Water permeability test (DIN 10480), initial surface absorption test (BS 1881) and rapid chloride permeability test (ASTM C1202 or AASHTO T 277) are suitable. Change in the mindset is required to accept pozzolanic additives to concrete or blended cements made with them.

ACKNOWLEDGMENT

We express our deepest gratitude to Prof. P.B. Sharma. for providing us with an environment to complete our project successfully. We are deeply indebted to our Head of the Department, Prof A K Gupta, who modeled us both technically and morally for achieving greater success in all our endeavors, we also thank all the staff members of our college and professors for their help in making this project a successful one.

Finally, we take this opportunity to extend our deep appreciation to our family and friends, for all that they meant to us during the crucial times of the completion of our project.

REFERENCES

[1] M.L.Gambhir , *Concrete Technology*, 4th edition, Tata McGraw Hill Education Private Limited, 7 West Patel Nagar, New Delhi 110008

[2] S.C. Rangwala, K.S. Rangwala, P.S. Rangwala, *Engineering Materials*, 38th edition, Charotar Publishing House Pvt. Ltd., ANAND 388 001 , Gujarat, India

[3] A.M. Neville, J.J. Brooks, *Concrete Technology*, 2nd edition, Pearson Publication.India

[4] Josef Hegger, Peter Doinghaus, *High Performance Steel and High Performance Concrete in Composite structures.*Technical Paper.

[5] Edward G. Nawy, *Fundamentals of High Performance Concrete*, 2nd edition, John Wiley & Sons, Canada..

[6] A.M. Neville, *Properties of Concrete*, 5th edition, Prentice Hall , Publications

[7] Mehta, *Concrete Microstructure Properties and Materials*, 3rd edition McGraw Hill Education Pvt. Ltd., India

Inter Project Validation for Change Proneness Prediction using Object Oriented Metrics

Dr. Ruchika Malhotra#, Megha Khanna*

#Department of Software Engineering, Delhi Technological University
Bawana Road, Delhi- 110042, India
ruchikamalhotra2004@yahoo.com
**AcharyaNarendraDev College*
University of Delhi, Govindpuri, Kalkaji, New Delhi- 110019, India
meghakhanna86@gmail.com

Abstract- **Developing prediction models for determining change prone classes of a software is a significant upcoming research area. Such models help usin effective and efficient resource utilization during maintenance phase. A crucial step in developing these prediction models is the use of training data of that project. Thus training data of a project dictates its prediction model. This dependency leads to difficulties in applying the prediction model to other projects. The training data of a project, consists of metrics data as well as change statistics. Although computation of metrics data is easy, but collection of change statistics is complex. In this paper we attempt to reuse the generated prediction model of one project and validate it on another project. For the purpose of our evaluation, we have used two open source projects written in Java language. The performance of the predicted models was evaluated using Receiver Operating Characteristic (ROC) analysis. The results of our study indicate that we can successfully apply the training sets for inter project validation. These results help us in optimizing time and effort required to generate training set of each project. It leads to efficient utilization of constraint resources and time.**

Keywords- **Change Proneness, Empirical validation,Inter Project validation,Object Oriented Metrics, Open Source, Receiver Operating Characteristic (ROC) analysis.**

1. INTRODUCTION

The principal challenge faced by software industry today, is efficient utilization of limited resources like time, cost and effort. Over the years, various researchers have conducted extensive studies to interrogate the relationship between object oriented (OO) metrics and software quality attributes like fault proneness and maintainability [1-5]. These studies assist us in effective and competent utilization of limited resources.

In order to establish and uncover the relationship between software quality attributes and OO metrics, we need prediction methods. An important step in developing prediction models is the learning process. With the help of learning, a prediction model adapts and adjusts its performance based on empirical data. The empirical data needed for learning process is called the training set. It helps in ascertaining possible predictive relationships between the dependent and independent variables. A training set consists of an input vector and an answer vector, which is used with a supervised learning method to train a knowledge database [6]. The training set is often used in alliance with a test set,which is used to examine the stability and fitness of the predictive relationship [6].

The training set for prediction model of a particular software, incorporates in it various characteristics of a software like its domain, environment, language etc [7]. Thus, a prediction model is highly dependent on its training set. However, possible reuse of training set of a particular project on another project could be of great advantage, as computation of training set involves a lot of effort. Such reuse of training set of a project to validate another project is called inter project validation. Our study aims at reusing training data for change proneness prediction for inter project validation.

Change proneness, i.e. the probability that a particular part of the software would undergo change in future, is a significant software quality attribute [8-10]. If a class is predicted as change prone, we can track its changes and test it rigorously leading to better quality software. Prediction of change prone classesusing OO metrics leads to effective resource utilization during maintenance [10].

The paper investigates the following research goals: (1) How accurately and precisely can predicted results of one project be applied on another project i.e. how successful is inter project validation? (2) What is the relationship between OO metrics and change proneness? (3) Which metrics are effective indicators of change proneness? (4) What is the accuracy of machine learning methods?
For the purpose of our analysis, we have taken two open

source software (Frinika and FreeMind) written in Java language. Two versions of both the software were taken and analyzed for changes. The changes were counted in terms of added, deleted and modified lines in the recent version classes with respect to previous version classes. OO metrics were also generated for all the classes of both the software. OO metrics and change statistics combined to yield data points. These data points form the training data of a project.

In order to evaluate inter project validation, we used Frinika software as our training set and validated it on Freemind software (test set). Apart from this, we also evaluated the results of Freemind using ten-fold cross validation.Both results were analyzed and yielded comparable results. The results were analyzed using Receiver Operating Characteristic (ROC) analysis. Though inter project validation for fault proneness has been investigated earlier [7], but to the best of our knowledge, no previous research has been conducted that validates inter-project prediction using change proneness data.

The main findings of our work are(1) Metrics CBO, RFC and LDC (refer Table 1) are important indicators for prediction of change prone classes (2) inter-project reuse of training data is possible for prediction of change prone classes using OO metrics and yields comparable results(3) the predicted results would help in efficient planning of testing resources during maintenance.
This paper is organized as follows, Section 2 summarizes related work. Section 3 explains the independent and the dependent variable. Section 4 specifies empirical data collection method. Section 5 enumerates the research methodology. The results of the study are presented in Section 6. Section 7 states threats to validity while Section 8 presents conclusion of the work.

2. RELATED WORK

Change proneness is a critical software quality attribute [8-10]. Han et al. [11]contributed in improvement of quality of design. TheyevaluatedJfreeChart, an open source project by applying Behavioral Dependency measurement (BDM). BDM is an approach to rate classes according to their probability of change.They concluded that BDM is a significant criteria for prediction of change proneness.
According to Ambros et al. [12] change coupling is defined as the dependency of various modules of a software on one another as they change and transform together. They analyzed the relationship between change coupling and software defects using correlation and regression analysis. They validated their study on three large software systems and found change coupling to be correlated more with large defects as compared to minor ones.

Sharafat et al. [9] analyzed the change history and source code of a class to predict the probability of change in the class. Reverse engineering techniques were used to analyze source code over various software releases and to collect code metrics. They also investigated the probability that a change would propagate a change in another class by studying the dependencies obtained from UML diagrams of the system.JFlex, a lexical analyzer generator for Java was used by them for evaluating their model.

Zhou et al. [8] studied Eclipse, a software developed in java programming language by analyzing three size metrics SLOC (Source lines of Code), NMIMP (Number of Methods Implemented in a class) and NumPara (sum of the number of parameters of the methods implemented in a class) to find out the confounding effect of class size on the relationship between OO metrics and change proneness.A confounding variable is the one which leads to distortion of the actual relationship between the dependent and the independent variable. They concluded that confounding effect of class size exists on the relationship between OO metrics and change proneness.

Watanabe et al. [7] have tried to evaluate the reuse of a fault prediction model on a C++ and Java open sourceproject. They also suggested compensation techniques to support inter language reuse of prediction models. They demonstrated inter language as well as inter project reuse of training data to cut costs and increase accuracy of results.
Zimmermann et al. [13] studied cross-project defect prediction on a huge scale by analyzing 12 real-world applications for 622 cross-project predictions. They concluded that cross project prediction using models from the projects in the same domain or with the same process are not always successful. They investigated and determined factors that influence cross –project predictions. They also derived decision trees to scrutinize the precision, recall and accuracy before attempting a prediction.

Malhotra et al. [10] investigated the relationship between OO metrics and change proneness of a class. For the purpose of their analysis they chose three open source software written in java language and used statistical as well as machine learning methods for predicting change proneness using ROC (Receiver Operating Characteristics) model. They concluded that relationship between OO metrics and change proneness of a class exists and machine learning methods outperformed statistical methods in this evaluation.

Lu et al. [14] examined 102 java systems by using statistical meta-analysis technique to ascertain the ability of 62 OO metrics to predict change proneness. They concluded

that size metrics exhibit moderate, coupling and cohesion exhibit lower than size metrics and inheritance exhibits poor capability to predict whether a class is change prone or not. They also ascertained using sensitivity analysis that their results are not biased on data selection.

3. INDEPENDENT AND DEPENDENT VARIABLE

In this section, we explain the independent and dependent variable in this study.

3.1 Independent Variable

The independent variable i.e. various OO metrics collected for each class of the software are summarized in Table 1. Several metrics are used in our study to account for various characteristics of software like size, coupling, cohesion, inheritance etc. The value for these OO metrics are obtained using Understand for Java (http://www.scitools.com/) software.

TABLE 1
SOFTWARE METRICS

S.No	Metric	Definition	Source
1.	Coupling Between Objects (CBO)	CBO for a class is a count of the number of other classes to which it is coupled and vice versa	[15, 18]
2.	Number of Children (NOC)	The NOC is the number of immediate subclasses of a class in a hierarchy.	[15, 18]
3.	Number of Methods per Class (NOM)	NOM is the total number of methods defined in the class.	[16, 18]
4.	Number of Attributes per Class (NOA)	NOA is the total number of attributes/variables defined in the class	[16, 18]
5.	Number of Instance Method (NIM)	It is the total number of Instance Methods.	[17]
6.	Number of Instance Variable (NIV)	It measures the relations of a class with other objects of the program.	[17]
7.	Number of Local Methods (NLM)	Number of local (not inherited) methods.	[19]
8.	Response For a Class (RFC)	A count of methods implemented within a class and the number of methods accessible to an object class due to inheritance.	[15, 18]
9.	Number of Local Default Visibility Methods (NLDM)	Number of local default visibility methods.	[19]
10.	Number of Private methods (NPRM)	Number of local (not inherited) private methods.	[19]
11.	Number of Protected Methods (NPROM)	Number of local protected methods.	[19]
12.	Number of Public Methods (NPM)	Number of local (not inherited) public methods.	[17]
13.	Number Of Lines (NL)	Number of all lines.	[19]
14.	Blank Lines Of Code (BLOC)	Number of blank lines of code.	[19]
15.	Source Lines Of Code (SLOC)	The number of lines that contain source code.	[19]
16.	Lines of Declarative Code (LDC)	Number of lines containing declarative source code.	[19]
17.	Lines of Executable Code (LEC)	Number of lines containing executable source code.	[19]
18.	Lines of Comment (LC)	Number of lines containing comment.	[19]
19.	Statement Count (SC)	Total number of declarative and executable statements.	[19]
20.	Depth of Inheritance (DIT)	The depth of a class within the inheritance hierarchy is the maximum number of steps from the class node to the root of the tree and is measured by the number of ancestor classes.	[15, 18]
21.	Lack of Cohesion (LC OM)	For each data field in a class, the percentage of the methods in the class using that data field; the percentages are averaged the subtracted from 100%.	[15, 18]
22.	Weighted Methods Per Class (WMC)	The WMC is a count of the sum of complexities of all methods in a class.	[15, 18]

3.2 Dependent Variable

Change in any software is necessary to keep it updated and useful. Estimation of change prone classes of a software could help in better enhancements to a software at lower costs [10]. The dependent variable in our study is change proneness. Change proneness of a class is the likelihood of change after the software has been delivered [10]. Change is the modification made in a class in terms of SLOC added, deleted or modified.

4. EMPIRICAL DATA COLLECTION

In this section we give introduction to our data sources and provide a description of our data collection method.

We examined two open source software written in java language. The source code of these open source software is available at http://sourceforge.net/. The change in a software is accounted for, by gathering changes in all classes of that particular software. Two versions of a particular software were taken and the change in classes present in both the versions was analyzed. The software we studied were a) Frinika, a complete music workstation software for Linux, Windows, Mac OSX and other operating systems running Java 1.5 b) Freemind, a mind mapper and hierarchical editor. The details of the software are provided in Table 2. It shows the version, release date, SLOC and number of classes for each of the software analyzed in this study.

TABLE 2
SOFTWARE DETAILS

Name	Version	Release Date	SLOC	No. of Classes
Frinika	0.2.0	04-07-2006	49,532 SLOC	539
	0.6.0	22-10-2009	1,13,093 SLOC	1273
Freemind	0.9.0 RC1	21-12-2008	83,495 SLOC	920
	0.9.0 RC7	06-03-2010	85,766 SLOC	955

4.1 Data Collection Method

The procedure adopted for collecting OO metrics and changes between multiple versions of a class is as follows. It is similar to the one followed by Malhotra et al [10]. First, we compute all OO metrics specified in Table 1 for initial versions (Frinika-0.2.0 and FreeMind 0.9.0RC1) with the help of Understand for Java (http://www.sitools.com/) software. These metrics are only computed for all classes of the software. The second step is to prepare comparable versions of the software. For this, common classes are extracted from both initial and recent versions of Frinikaand FreeMind software. These classes are compared line by line with the help of Understand for Java to compute number of SLOC ADDED, DELETED and MODIFIED. TOTAL CHANGE for class is calculated by counting each ADDED or DELETED line as one SLOC change and each MODIFIED line as two SLOC changei.e. one deletion followed by one addition [8,10]. Another binary variable ALTER is defined which has the value ‘yes’ if TOTAL CHANGE > 0 or ‘no’ otherwise [10]. The change statistics (ADDED SLOC, DELETED SLOC, MODIFIED SLOC, TOTAL CHANGE and ALTER) and OO metrics are put together to generate data points. Malhotraet al. [10] provide a detailed and step by step procedure for data collection.

ForFrinika, out of 539 classes of Frinika 0.2.0, 80 classes were those which were deleted i.e. not present in 0.6.0. Only 248 classes were present in both the versions leading to 248 data points for Frininka. After data analysis, it was found that out of 248 data points 127 (51%) classes were altered while 121 classes were not changed in version 0.6.0. For FreeMind, all 920 classes were present in both the versions but library classes and unknown classes were discarded. This resulted in 657 data points.Out of 657 data points, 68 (10%) classes were found to be changed while the other 589 classes were not changed in 0.9.0 RC7 version.

5. RESEARCH METHODOLOGY

This section explains the various Machine Leaning (ML) methods we use for our analysis. We have used the default setting of WEKA tool.

5.1 Correlation based Feature Selection

While using a machine learning method, it is essential to diagnose and observe representative set of features [19]. Features in our case would be various OO metrics. A well established feature set is the one in which all features are highly correlated with the dependent variable, but these features should not be dependent on one another [20]. The dependent variable for our study is change prone classes.

Correlation based Feature Selection (CFS) aids in identifying repetitive and unwanted noisy features which should be removed. It has been established that classification results are more accurate if we use a reduced feature set as compared to a complete feature set [20]. CFS uses a correlation method among variables to identify good as well as noisy features. With the help of CFS we reduce execution time and improve predictive accuracy [21].

5.2 Decision Table

A Decision Table (DT) is a hierarchical structure of data in which complex data entries are present at higher levels of the table [22]. These complex entries are broken down into simpler entries with the help of additional attributes to form the table hierarchy. Attribute selection for break down of the table is measured by performing cross validation of all combinations of present attributes [6]. The subset which gives the best result is chosen.

5.3 Bayesian Network

A Bayesian Network (BN) is an interconnected network of nodes, where each node represents a random variable and all directed edges connecting these nodes represent probabilistic dependencies among nodes [6]. A BN should not contain a cycle. BN helps in computing joint probability distribution among a set of random variables. The structure of a BN represents conditional independence assertions [23]. BN can easily handle incomplete data sets and also allows us to investigate casual relationships.

5.4 Adaptive Boosting

Adaptive Boosting (AB) is a technique in which we combine many weak performance classifiers to obtain a strong performing classifier. This is done by building subsequent models which are based on gaining expertise on instances which are incorrectly classified by previous models [6]. It uses voting and assigns weights to its instances on the basis of performance [6]. The better performing model is given more weight. AB is sensitive to noisy data and outliers.

6. ANALYSIS RESULTS

In this section we describe the analysis performed for inter-project validation for change proneness using object oriented metrics. We employed DT, BN and AB machine learning methods for our analysis. The models predicted were applied to all the data points collected from the software. The following measures are used to evaluate the performance of each predicted model:

1. The sensitivity and specificity of the model are calculated to estimate correctness of the model. The percentage of classes correctly predicted to be change prone is known as sensitivity of the model [10]. The percentage of classes predicted not to be change prone is called specificity of the model. Ideally, both sensitivity and specificity should be high to predict change-prone and non change-prone classes [10].

2. Receiver operating Characteristic (ROC) analysis: The performance of the outputs, of the predicted models were assessed using ROC analysis. The ROC curve is defined as a plot of sensitivity on the y-coordinate versus 1-specificity on the x-coordinate [5]. An optimal choice of the cutoff point (that maximizes both sensitivity and specificity) can be chosen from the ROC curve.Area Under the ROC Curve(AUC) is a combined measure of sensitivity and specificity [5,10]. In order to compute theaccuracy of the predicted models, we use the area under the ROC curve.

3. Ten-fold cross validation of all prediction models is performed which analyzes the accuracy of the model. The data set is randomly divided into ten subsets. A number of iterations are performed by using one of the ten subsets as the test set while the other nine subsets are used for training. Therefore, we get the change proneness prediction for all the ten subsets [24].

6.1 Descriptive Statistics

A table for each software, Frinika and FreeMindis presented in the following subsection which show Minimum (Min), Maximum (Max), Mean (Mean), Median (Med) and Standard Deviation (SD) for all metrics considered in this study. The following observations are made from Tables 3 and 4.

1. The size of class measured in terms of SLOC ranges from 3 to 1538 (Frinika) and 1 to 1513 (FreeMind).

2. The mean values of NOC (Frinika -- 0.53, FreeMind -- 0.32) and DIT (Finika -- 3.32, FreeMind -- 1.98) are poor for both the projects which shows that there are only few subclasses and inheritance is not much used in these projects; similar results have been shown by others [17, 25-26].

3. The LCOM measure has high values (upto 100) in both data sets.

TABLE 3
DESCRIPTIVE STATISTICS FOR FRINIKA SOFTWARE

Metric	Min	Max	Mean	Med	SD
CBO	0	32	7.62	2	61.78
NOC	0	12	0.53	0	16.33
NOM	0	20	0.97	0	17.62
NOA	0	28	2.54	1	25.49
NIM	0	79	13.84	3	109.98
NIV	0	135	9.25	2	75.07
NLM	0	79	14.81	3	117.53
RFC	0	120	23.94	5	189.02
NLDM	0	9	0.79	0	16.95
NPRM	0	15	0.75	0	16.88
NPROM	0	10	0.36	0	16.03
NPM	0	73	12.9	3	102.75
NL	4	2047	255.94	57	2018.52
BLOC	0	264	44.64	10	351.89
SLOC	3	1538	167.15	40	1319.53
LDC	2	365	47.35	12	373.38
LEC	0	1054	85.72	18	679.89
LC	0	381	49.36	5	390.68
SC	2	1270	120.38	28	951.91
DIT	1	4	3.32	2	30.35
LCOM	0	100	93.64	50	735.27
WMC	0	163	28.25	6	222.92

TABLE 4
DESCRIPTIVE STATISTICS FOR FREEMIND SOFTWARE

Metric	Min	Max	Mean	Med	SD
CBO	0	148	4.96	3	7.87
NOC	0	18	0.32	0	1.49
NOM	0	63	0.3	0	2.63
NOA	0	37	0.91	0	2.81
NIM	0	129	7.39	4	12.66
NIV	0	105	2.38	1	5.94
NLM	0	129	7.69	4	12.97
RFC	0	222	14.83	6	23.83
NLDM	0	21	0.33	0	1.22
NPRM	0	16	0.95	0	2.15
NPROM	0	23	0.47	0	1.75
NPM	0	116	5.95	3	11.06
NL	1	2829	109.07	46	217.97
BLOC	0	244	13.5	6	27.42
SLOC	1	1513	78.75	33	143.15

LDC	1	471	23.64	11	41.84
LEC	0	766	38.46	15	75.84
LC	0	1544	17.86	3	70.52
SC	1	1152	56.07	24	103.51
DIT	1	6	1.98	2	1.04
LCOM	0	100	34.88	33	35.97
WMC	0	250	14.26	7	26.69

TABLE 5
DESCRIPTIVE STATISTICS OF DEPENDENT VARIABLE

Software Name	Dependent variable	Min	Max	Mean	SD
Frinika	TOTAL CHANGE	0	1166	46.202	140.024
FreeMind	TOTAL CHANGE	0	94	1.516	8.358

Table 5 presents the descriptive statistics of Frinika and FreeMind software. It displays the TOTAL CHANGE, the dependent variable which was calculated (Section 4.1) for both the software with its Minimum (Min), Maximum (Max), Mean and Standard Deviation (SD).

6.2 Result Analysis for FreeMind Software using Ten-Fold Cross Validation

First we applied CFS on data points generated by FreeMind software to select metrics for predicting change proneness. According to Table 6, the metrics selected were CBO, NOM, NOA, RFC, NLDM, NPRM, NPROM, NPM, BLOC, LDC, LC AND SC. Table 7 shows the validation results on FreeMind software using ten-fold cross validation. DT in this case gives the best results with AUC of 0.733 and specificity and sensitivity of 68.3% and 69.1% respectively. The cut off point for DT is 0.097. AB also gives good results with AUC of 0.709 and specificity and sensitivity of 60.6% and 69.1% respectively. The cut off point for AB is 0.061. The results of BN are comparable. BN gives AUC of 0.673. The specificity, sensitivity and cut off point for BN is 63.7%, 63.2% and 0.075 respectively.

TABLE 6
METRICS SELECTED BY CO-RELATION BASED FEATURE SELECTION

Software Name	Metrics Selected
Frinika	CBO, RFC, NIV, BLOC, LDC
FreeMind	CBO, NOM, NOA, RFC, NLDM, NPRM, NPROM, NPM, BLOC, LDC, LC, SC

TABLE 7
FREEMIND VALIDATION RESULTS USING TEN-FOLD CROSS VALIDATION

Method	Specificity	Sensitivity	Cutoff Point	AUC
DT	68.3	69.1	0.097	0.733
BN	63.7	63.2	0.075	0.673
AB	60.6	69.1	0.061	0.709

6.3 Result Analysis of model prediction using Frinika as Training Set

For using Frinika as training set and FreeMind as test set, we first selected metrics that predict change proneness by applying CFS. CFS was applied on data points generated by Frinika software. The metrics selected after applying CFS on Frinika were CBO, RFC, NIV, BLOC and LDC (Table 6). After applying CFS on Frinika data set, we use it as training set and supply FreeMind data points as test data using WEKA tool. The validation results obtained are present in Table 8. Here AB gives the best results with AUC of 0.767 and8 specificity and sensitivity of 69.9% and 70.6% respectively. The cut off point for AB is 0.742. After AB, DT showed good results with AUC of 0.742. The specificity and sensitivity with DT was 62.3% and 77.9% respectively and the cut off point was 0.608. Alternatively, the results with BN are also good with AUC of 0.727, specificity and sensitivity of 63.0% and 72.1% respectively and cut off point of 0.653. 6.4 Discussion of Results

TABLE 8
VALIDATION RESULTS USING FRINIKA AS TRAINING SET

Method	Specificity	Sensitivity	Cutoff Point	AUC
DT	62.3	77.9	0.608	0.742
BN	63.0	72.1	0.653	0.727
AB	69.9	70.6	0.613	0.767

The data points generated by FreeMind were validated using a) ten-fold cross validation and b) by using Frinika data set as the training set. FreeMind validation results for predicting change prone classes using both the above mentioned approaches were comparable. All the three machine learning techniques DT, BN and AB gave competent results for predicting change proneness. While DT, BN and AB gave AUC value of 0.733, 0.673 and 0.709 respectively when we use ten-fold cross validation, the AUC values while using Frinika as the training set were 0.742 0.727 and 0.767 respectively for DT, BN and AB.These results indicate that we can use a training set of one software project on anothersoftware project effectively while applying prediction models for change proneness. Such a technique is called inter-project validation. We concluded successful inter-project validation as results using FreeMind's own training set in ten-fold cross validation and the results by using Frinika (a different software) as the training set were comparable and competent.Reuse of training set for other software projects helps in better utilization of limited resource like time, cost and effort. Computing training sets specific to each software is a laborious and time consuming process. Thus effective re-utilization of already computed data sets can be done by employing them as training sets while applying prediction models on new data sets (software). Figure 1 below shows the process of inter-project validation using FreeMind as test set and Frinika as training set.

Another result obtained is that CBO, RFC and LDC metrics are good indicators of change proneness. These metrics were selected in the feature sets of both the data sets, after applying CFS.Prediction of change prone classes of a software,help us in better planning and administrating resources during maintenance. Figure 2 below shows the ROC curves obtained while applying DT, BN and AB on FreeMind using ten-fold cross validation while Figure 3 shows ROC curves on FreeMind, when Frinika data set is used for training.

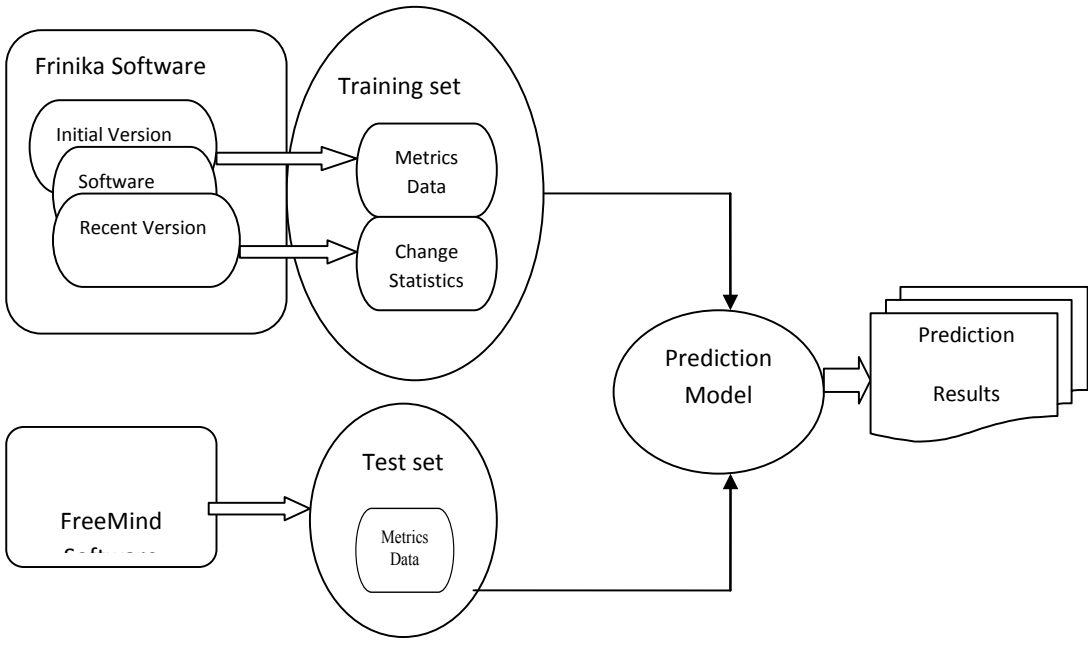
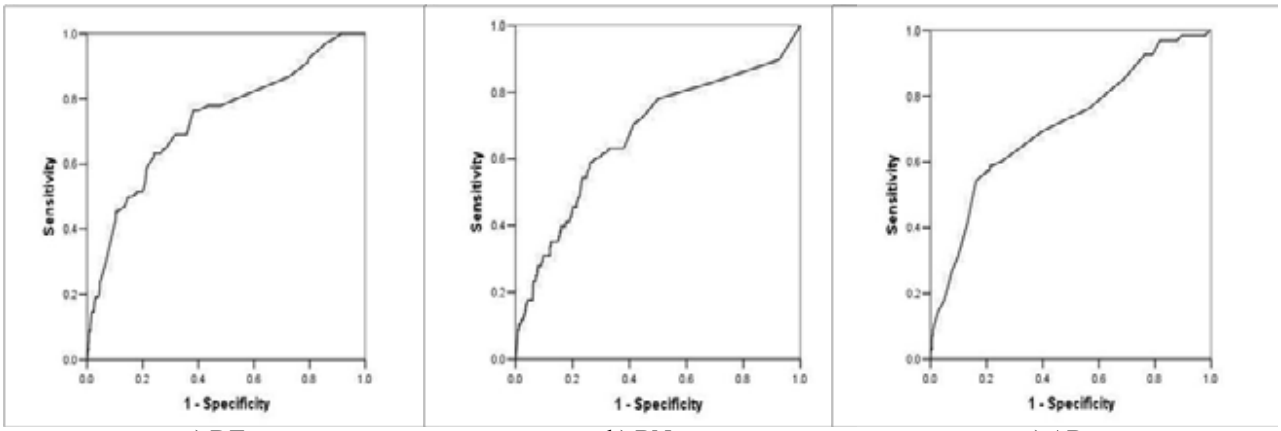
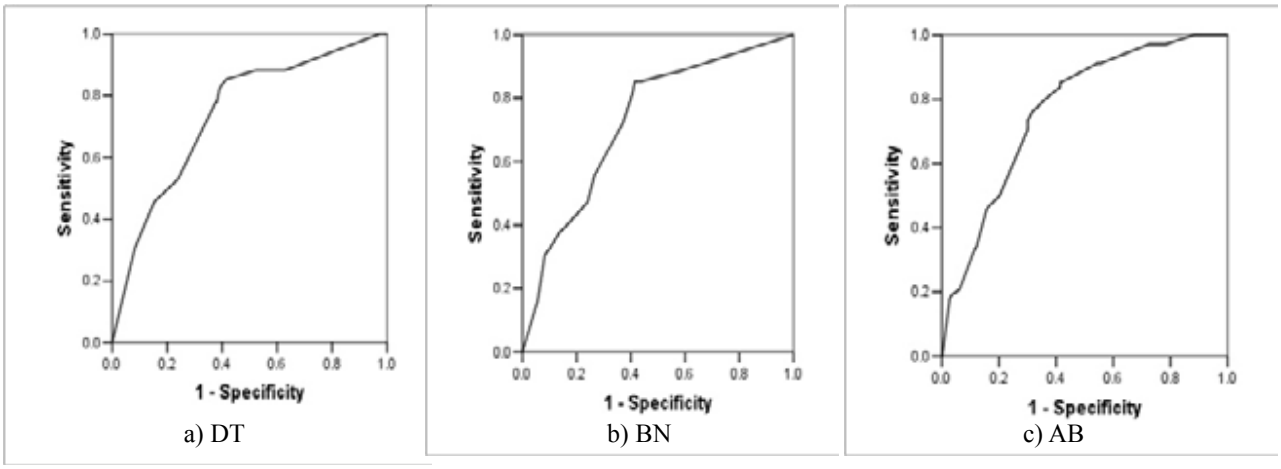


FIGURE : INTER PROJECT VALIDATION DIAGRAM



a) DT b) BN c) AB

FIGURE 2 : ROC CURVE FOR FREEMIND SOFTWRAE USING TEN-FOLD CROSS VALIDATION a) DT b) BN c) AB



a) DT b) BN c) AB

FIGURE 3 : ROC CURVE FOR FREEMIND SOFTWARE USING FRINIKA AS TRAINING SET a) DT b) BN c) AB

7. THREATS TO VALIDITY

In this section we discuss the various threats to validity: construct, external and internal which are prevalent in our study.

7.1 Construct Validity

The dependent and independent variables are measures of certain concepts. Construct validity is attributed to the degree of accuracy to which these concepts are measured with these variables [8]. Our independent variable, i.e. the various OO metrics (Table 1) have already been validated in previous studies[25, 27-28] as to whether they accurately represent the concepts like coupling, cohesion, inheritance etc. The values of various metrics are collected using Understand for Java tool. The tool does not specify the method followed for collection of these metrics for individual classes which would be a significant indicator of their construct validity. Change proneness, our dependent variable is collected manually by setting the ALTER variable in accordance with the change in the corresponding class. This eliminates construct validity threat with respect to the dependent variable. Software changes can be categorized as corrective, adaptive, perfective or preventive [8, 10, 29]. For our study, we did not classify the various kinds of changes as in [10].

7.2 External Validity

External validity is associated with the extent of generalization of results which can be demonstrated by comparative analysis on various data sets [8,10]. We cannot assure whether our results are universally applicable as we have not accounted for different programming languages and environments. This threat can only be minimized by conducting replicated studies with data sets having different characteristics.

7.3 Internal Validity

Internal validity is the extent to which we can summarize the casual effect of independent variable like coupling, cohesion, inheritance etc on the dependent variable i.e. change proneness [8]. In order to comment on casual effect we need to perform controlled experiments where a particular OO construct (like coupling) is varied and other constructs like cohesion and inheritance remain constant [2]. It is difficult to perform such an experiment and this threat exists in our study.

7. CONCLUSION AND FUTURE WORK

The goal of our research was to empirically examine inter project validation for predicting change prone classes using OO metrics. Based on the studies of data sets obtained from two open source software Frinika and FreeMind, we analyzed the performance of predicted models using ROC analysis. We applied machine learning methods to predict the effect of OO metrics on change proneness. We also took account of changes in the classes while predicting change proneness. Our main results are summarized as follows.

1. The metrics CBO, RFC and LDC are significant indicators of predicting change prone classes as in both the data sets, these metrics were chosen after applying feature subset selection method CFS. A good feature sets has features that are highly correlated with the dependent variable but are uncorrelated with each other [19]. CFS uses a correlation method among variables to identify good as well as noisy features.

2. The predicted results indicate that inter project validation of OO metrics for change proneness prediction is valid and yields competent results.While DT, BN and AB gave AUC value of 0.733, 0.673 and 0.709 respectively when we use ten-cross validation, the AUC values while using Frinika as the training set were 0.742 0.727 and 0.767 respectively for DT, BN and AB. These results indicate that we can use a training set of one software project on another software project effectively while applying prediction models for change proneness.While using ten-fold cross validation we use the training set of the same project which incorporates various software characteristics like domain, language etc of the particular software. But when we used Frinika as the training set and applied on FreeMind as the test set i.e. reuse a training set of a different project on another, the results were still comparable to that often-fold cross validation.This result helps research analysts and practitioners in reusing prediction models based on one software project on another project for prediction of change prone classes. This leads to effective and efficient utilization of resources and effort and time reduction as we do not need to compute training sets of specific software but can reuse an already existing set of a different software.

3. The results help us in effective utilization of constraint testing and maintenance resources. Change prone classes of a software need to be tested exhaustively during maintenance. Thus more testing resources can be assigned to these change prone classes for rigorous testing. Meticulous testing of change prone classes leads to a good quality software. Our results indicate that CBO, RFC and LDC are suggestive indicators of change prone classes. We can plan our limited testing resources in such a manner that they give competent good quality software products.

The results of our study are valid for object oriented medium systems. We plan to replicate our studies on data sets having different characteristics such as datasets with different programming languages and environments. We also plan to use other machine learning algorithms like genetic algorithms for inter-project analysis.

REFERENCES

[1] W. Li and S. Henry, "Object Oriented Metrics that Predict Maintainability,*"Journal of Systems and Software,* vol. 23, pp. 111-122,1993.

[2] L. Briand, J. Wust and H. Lounis,"Replicated Case Studies for Investigating Quality Factors in Object Oriented Designs.*"Empirical Software Engineering: An International Journal,* vol 6, pp. 11-58, 2001.

[3] V.R. Basili, L.C. Briand and W.L. Melo, "A Validation of Object-Oriented Design Metrics as Quality Indicators," IEEE *Transactions on Software Engineering,* vol. 22, no. 10, pp. 751-761, 1996.

[4] K.K. Aggarwal, Y. Singh, A. Kaur and R. Malhotra,"Empirical Analysis for Investigating the Effect of Object-Oriented Metrics on Fault Proneness: A Replicated Case Study.,"Software Process: Improvement and Practice, vol. 16, no.1, pp. 39-62, 2009.

[5] Y. Singh, A. Kaur and R. Malhotra. "Empirical Validation of Object-Oriented Metrics for Predicting Fault Proneness,*" Software Quality Journal,* vol. 18, no.1, pp. 3-35, 2010.

[6] I.H Witten and E. Frank, D*ata Mining : Practical Machine Learning Tools and Techniques,* 2nded, Elsevier, 2005.

[7] S. Watanbe, H. Kaiya and K. Kaijiri, "Adapting a Fault Prediction Model to Allow Inter Language Reuse,*" PROMISE 2008*, 2008.

[8] Y. Zhou, H. Leung and B. Xu,"Examining the Potentially Confounding Effect of Class Size on the Associations between Object Oriented Metrics and Change proneness," *IEEE Transactions on Software Engineering,* 35, no. 5, 2009

[9] A.R. Sharafat and L. Tavildari, "Change Prediction in object Oriented Software Systems : A Probabilistic Approach," *11th European Conference on Software Maintenance and Reengineering,* 2007.

[10] R. Malhotra and M. Khanna, "Investigation of Relationship between Object-oriented Metrics and Change Proneness," *International Journal of Machine Learning and Cybernetics,* Springer ,2012

[11] A.R. Han, S. Jeon, D. Bae and J. Hong,"Behavioral Dependency Measurement for Change Proneness prediction in UML 2.0 Design Models," *Computer Software and Applications 32nd Annual IEEE International,* 2008.

[12] M. D'Ambros, M. Lanza and R. Robbes, "On the Relationship Between Change Coupling and Software Defects," *16th Working Conference on Reverse Engineering,* pp. 135-144, 2009.

[13] T. Zimmermann, N. Nagappan, H. Gall, E. Giger and B. Murphy, "Cross-project Defect Prediction A Large Scale Experiment on Data vs. Domain vs. Process", in Proceedings ofthe 7th joint meeting of the European Software EngineeringConference and the ACM SIGSOFT International Symposiumon Foundations of Software Engineering. ACM, 2009, pp.91–100.

[14] H. Lu, Y. Zhou, B. Xu, H. Leung and L. Chen, "The ability of object-oriented metrics to predict change-proneness: a meta-analysis",*Empirical Software Engineering Journal, vol.* 17, no.3, pp 200-242, 2012

[15] S.R. Chidamber and C.F. Kemerer,"A Metrics Suite for Object Oriented Design," *IEEE Transactions of Software Engineering,* vol. 20, no.6, pp. 476-493, 1994.

[16] B.Henderson-sellers. Object Oriented Metrics. Measures of Complexity. Prentice Hall

[17] M. Lorenz and J. Kidd, "Object-Oriented Software Metrics,"*Prentice Hall Object-Oriented Series,* Englewood Cliffs, N.J., 1994.

[18] K.K. Aggarwal, Y. Singh, A. Kaur and R. Malhotra,"Empirical Study of object-oriented metrics",*Journal of Object Technology,* 5, no.8, pp. 149-173, 2006.

[19] http://www.scitools.com/features/metrics.php

[20] M.A. Hall,"Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning," *Proceeding of the Seventeenth International Confrence on Machine Learning*,. pp. 359-366, 2000.

[21] K. Michalak and H. Kwasnicka,"Correlation-based Feature Selection Strategy in Neural Classification," *Sixth International Conference on Intelligent Systems Design and Applications*, Vol 1, pp. 741-746.

[22] R. Kohavi and D. Sommerfield, "Targeting Business Users with Decision Table Classifiers," *Proceedings of IEEE Symposiomon Information Visulaization*, 1998.

[23] F. Ruggeri , F. Faltin and R. Kenett,En*cyclopedia of Statistics in Quality & Reliability*, Wiley & Sons (2007).

[24] M. Stone,"Cross-validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society Series A*, Vol. 36, pp.111-14, 1974.

[25] L. Briand, J. Wust, J. Daly and D.V. Porter,"Exploring the Relationships between Design Measures and Software Quality in Object-oriented Systems," *Journal of Systems and Software,* 51, no.3, pp. 245-273, 2000.

[26] S.R. Chidamber, D.P. Darcy and C.F. Kemerer., "Managerial Use of Metrics for Object-Oriented Software: An Exploratory Analysis," *IEEE Transactions on Software Engineering*, 24, no.8, pp. 629-639, 1998.

[27] L. Briand, J. Daly, J. Wust, "A Unified Framework for Cohesion Measurement in Object Oriented Systems", *Empirical Software Engineering Journal*, 3, no.1, pp 65-117, 1998.

[28] L. Briand, J. Daly, J. Wust, "A Unified Framework for Coupling Measurement in Object Oriented Systems", *IEEE Transactions on Software Engineering*, 25, no.1, pp 91-121, 1999.

[29] K.K. Aggarwal, Y.Singh, *Software Engineering*, 3rd Edition, New Age International Publishers (2008).

ABOUT THE AUTHORS

Ruchika Malhotra. is an Assistant Professorat the Department of Software Engineering, Delhi Technological University (formerly Delhi College of Engineering), Delhi, India. She was an assistant professor at the University School of Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. She received her master's and doctorate degree in software engineering from University School of Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. Her research interests are in software testing, improving software quality, statistical and adaptive prediction models, software metrics, neural nets modeling, and the definition and validation of software metrics. She has published more than 70 research papers in international journals and conferences. Malhotra can be contacted by email at ruchikamalhotra2004@gmail.com

Megha Khanna. is currently working as an Assistant Professor in Acharya Narendra Dev College, University of Delhi. She completed her Masters degree in Software Engineering (2010) from University School of Information Technology, Guru Gobind Singh University, India. She received her graduation degree in Computer Science (Hons.) (2007) from Acharya Narendra Dev College, Delhi University, India. Her research interests are in improving software quality, neural nets, artificial intelligence and the definition and validation of software metrics. She can be contacted by e-mail at meghakhanna86@gmail.com.

Article Title Page

Modeling of Information Security Management Parameters in Indian Organizations using ISM and MICMAC Approach

Author Details:

Muktesh Chander
Department of Management Studies, Indian Institute of Technology, Delhi,
New Delhi, India

Sudhir K. Jain
Department of Management Studies, Indian Institute of Technology, Delhi,
New Delhi, India

Ravi Shankar
Department of Management Studies, Indian Institute of Technology, Delhi,
New Delhi, India

Corresponding author: Muktesh Chander
mukteshchander@gmail.com

Biographical Details (if applicable):

Muktesh Chander is a Graduate in Electronics and Communication Engineering from Delhi College of Engineering. He also holds a Law degree from Delhi University and MA in Criminology and Forensic Science from M. S. University, Tirunelveli. He is currently a Ph.D. scholar at Indian Institute of Technology, Delhi. Muktesh Chander is the corresponding author and can be contacted at: mukteshchander@gmail.com

Prof. Sudhir K. Jain is MHRD IPR Chair Professor and the Chairman of 'Economics & Entrepreneurship' group. He obtained B.Sc. and M.A. (Economics) degrees from Rajasthan University, Jaipur and Ph.D. from I.I.T. Kanpur. He has co-authored “Managerial Economics” with Petersen & Lewis, published by Pearson Education in 2005. Earlier, he headed National Institute for Entrepreneurship & Small Business Development (NIESBUD) under the Ministry of MSME as its Executive Director. Currently, he is Head of the Department of Management Studies, I.I.T. Delhi. He has published over 100 research papers in reputed journals and conference proceeding. He has also been associated with World Intellectual Property Organization (WIPO) in various capacities including Tutor in Distance Learning Program of its World Wide Academy.

Prof. Ravi Shankar is Professor of Operations and Information & Telecommunication Technology Management. He is Group Chair of Operations & Supply Chain Management and Sectoral Management groups at the Department of Management Studies at Indian Institute of Technology Delhi, India. His publications have appeared in various journals including the European Journal of Operational Research, International Journal of Production Research, Computers and Industrial Engineering, International Journal of Production Economics, Computers and Operations Research, International Journal of Supply Chain Management, etc. He has co-authored text books on Technology Management, Supply Chain management, Operations Management, Enterprise resource Planning, Total Quality

Management, etc. He is the Executive Editor of Journal of Advances in Management Research. He may be contacted at: ravi1@dms.iitd.ac.in

Structured Abstract

Purpose: Increasing use of computers, computer networks and communication technology has resulted in huge amount of vital information being exchanged, stored and processed on computers by individuals, various organizations, governments and their institutions. As the digital dependence of organizations is increasing at a very rapid pace, Information Security Management is becoming an important issue in modern corporate governance. The present research aims to provide a comprehensive framework for the various important factors of Information Security Management at the organizational level leading to a culture of information security in the organization. The purpose of this paper is to identify various information security management parameters and develop a conceptual framework for it.

Design/methodology/approach: Interpretive Structural Modeling (ISM) and MICMAC approaches have been used to identify and classify the key factors of information security management based on the direct and indirect relationship of these factors.

Findings: The research presents a classification of key parameters according to their driving power and dependence which enable information security management in an organization. It also suggests parameters on which management should pay more attention.

Research Limitations/Implications: In the present study, 12 parameters were identified based on literature study and expert help. It is possible to identify some more parameters for ISM development. Help of experts was also taken to identify the contextual relationship among the variables for ISM model. This may introduce some element of bias. Although a relationship model using ISM has been developed but it has not been validated statistically. For future research, it is suggested that structural equation modelling (SEM) technique may be used to corroborate the findings of ISM. Some of the variables have been clubbed together being a part of a subset due to their similar nature but it is possible to treat them as independent variables. Future researches may establish their interrelationships also.

Practical Implications: The paper has tremendous practical utility for organizations which want to reap the benefits of information and communication technology for their growth but are struggling to find a right approach to deal with information security breach incidents.

Originality/value: Development of a framework for information security management in an organization is the major contribution of this paper. This would be of help to strategic managers in managing information security with emphasis on key parameters identified here.
.

Keywords: Information Security, Information Security Management, Interpretive Structural Modelling (ISM), MICMAC, key variables.

Article Classification: Research Paper

For internal production use only

Running Heads:

Modelling of Information Security Management Parameters in Indian Organizations using ISM and MICMAC Approach

Muktesh Chander, S. K. Jain and Ravi Shankar

Department of Management Studies, Indian Institute of Technology, Delhi, New Delhi, India

1. Introduction

Information and Communication Technology (ICT) has percolated at almost all levels of human activities in a big way. Private companies, government institutions, organizations and citizens are using computers, computer networks and Internet like never before. World is fast moving towards a connected digital society and hence everyone need to be secured in the cyber space. Information Security stands on a tripod of confidentiality, integrity and availability of data and services commonly known as CIA triad (ISO/IEC 27001:2005). Without proper information security mechanism, computerization can never achieve its fullest potential as a serious business and governance tool. The critical sectors of a modern nation are heavily dependent on a complex network of interdependent computer systems. Smooth functioning of these sectors is important for national security also. Any vulnerability in the computer systems controlling these sectors can not only be devastating in terms of massive losses but can also threaten a nation. There is a need to adopt a holistic approach to information security as an organizational priority to secure their digital assets. Identification of key success factors/enablers of information security and their interrelationship in an organization is useful for strategic decision making in this regard.

2. Literature review

The status of information security in various organizations in most of the countries is far from satisfactory. Even in the most advanced countries organizations are victim of information security breaches. Information theft, loss or attack has emerged as the most commonly reported frauds globally (Kroll, 2010). According to Information Security Breaches Survey Report 2010 (PWC, 2010), 92% of large size respondent organizations and 83% of small ones experienced an information security breach incident in the last one year. According to another survey, 60 % organizations encountered one or more cyber security incidents in a year (CSO, 2010). Computer security breach incidents handled by Indian Computer Emergency Response Team (CERT-In) increased by 222% in 2009 as compared to last year (CERT, 2009). Cyber crimes are on the rise and sophistication of cyber criminals has been increasing. Hence the organizations are realizing the importance of information security measures and are prepared to spend resources for them.

2.1. Identification of information security enablers

Information security in an organization is dependent numerous factors. Starting from physical access control to technical solutions, information security has evolved as a multi-disciplinary subject with recent researches highlighting the importance of human and organizational factors. From the survey of existing

literature and research, following factors have been identified as critical in information security management.

2.1.1. Top management commitment

Top management is responsible for each and every activity at all levels of the organizations (Singh & Kant, 2008). Ultimate responsibility for managing information security is borne by corporate management which provides the resources and sets the requirements. On this basis the IT security manager promotes and coordinates security activities (Kajava *et al.*, 2006). Top management commitment is a prerequisite of any information security program. This shows a clear direction and support to information security efforts. Vaish and Verma (2010) have found senior management commitment as the most important parameter in information security management. ISO/IEC 27001(2005) mandates that management should actively support security within the organization through clear direction, demonstrated commitment, explicit assignment and acknowledgment of information security responsibilities. ISO/IEC 27002(2005) also highlights the role of management in information security management. Draft ISO Guide 83 (2011) mentions that top management must demonstrate its commitment. Björck(2001) has also identified top managements' commitment as one of the six important parameters of information security management in Swedish organizations. Similar finding were reported by Kankanhalli *et al.* (2003) in small and medium-sized Singaporean enterprises and by Sami (2006) in Jordan. Von Solms (2006) has argued that top management commitment is an important component of corporate security governance. Top management support is also essential for the allocation of resources (Avolio, 2000). Information security must be understood, supported and funded from the top down (Harris, 2010). Strong top management support is critical to information security management (NIST, 2008; Knapp *et al.,* 2006; Wright, 2008).

2.1.2. Identification and classification of information assets of organization

Anything which needs to be protected must be first identified. The amount of time, effort and money, which an organization wants to spend on protecting any information asset, would depend on its value, utility, sensitivity, criticality and legal requirements. ISO/IEC 27001:2005 and ISO/IEC 27002:2005 advocate identification of information assets, their classification and ownership. The identification of information processed on an information system is essential to the proper selection of security controls (NIST, 2008). Englesman (2007) has provided a framework for identification and valuation of information assets.

2.1.3. Technological tools and solutions

Technological tools and solutions are the foundation on which information security stands. Technology is one of the tools which cyber criminals use to commit cyber crimes and hence it can be countered using adequate counter technologies only. They include anti-virus, anti-spyware, firewalls, access control, intrusion detection and prevention systems, cryptography, backup tools, content management tools, patch management tools etc.(Harris, 2010; NIST, 2009; Belasco and Wan, 2006; Gallaher *et al.*, 2008). ISO/IEC 27039 currently under draft stage deals with selection, deployment and operations of Intrusion Detection [and Prevention] Systems (IDPS).

2.1.4. Providing information security policies and procedures

A policy is typically described as a set of basic principles and guidelines, formulated and enforced by top management of an organization in order to achieve a certain goal. Policy prescribes the aims, objectives and the targets that need to be used to achieve the goals (Sushil *et al*., 2006). Information Security Policy is an aggregate of directives, rules and practices that prescribe how an organization manages, protects and distributes information (NIST, 2006 A). Hong *et al*. (2006) define information security policy as the set of rules set-up for the use of information assets and the statement set-up for the security priorities to achieve the objectives of an organization.

2.1.5. Providing information security organizational structure and resources

Any information security program requires adequate organizational structure and resources. Top management commitment it is of no use unless it results in providing adequate manpower and other resources for the information security functions. Tudor (2001) has highlighted the importance of security organization and infrastructure. ISO/IEC 27001 (2005) provides for provision of resources to cover all the aspects of information security management. All information security projects need budgeted resources (Björck, 2001). Information security organizational structure should be appropriate for the organization it supports (NIST, 2006 A). Organizational resources are the fundamental requirement to enforce and monitor the implementation of information security (Al-Awadi, 2009). Draft ISO Guide 83 (2011) also refers to determining and providing resources. Draft ISO/IEC TR 27016 dealing with the organizational economics of information security management also highlights the importance of allocating adequate budget for information security functions.

2.1.6. Awareness, education and training of stakeholders

It is most crucial that employees, customers, vendors and all other stakeholders are fully made aware of the various threats to information security and their role and contribution in the overall goal of information security in an organization. Siponen (2000) describes information security awareness as a state where users in an organization are aware of their security mission. Several international standards such as ISO 27001:2005, Control Objectives for Information and related Technology (COBIT), Payment Card Industry-Data Security Standard (PCI-DSS) have highlighted the need of information security awareness and training. Awareness of the risks and available safeguards is the first line of defence for the security of information systems and networks (OECD, 2002). Information security awareness amongst users is of core importance (Von Solms and Von Solms, 2004). Protection of an organization begins by making sure that the employees understand their roles and responsibilities in safeguarding sensitive data and protecting company resources and assist the organization in keeping computers and network safe(ENISA, 2009). Raising awareness of cyber security at all levels of government is one of the priority areas identified in the Cyber Security Strategy of the United Kingdom (UK, 2009). Similar priority has been mentioned in the Australian Cyber Security Strategy with emphasis on conducting education and awareness raising activities to promote a culture of cyber security amongst all Australian ICT users (Australia, 2009). Estonian Cyber Security Strategy, (Estonia, 2008) has highlighted the need for raising awareness of information security among all computer users with particular focus on individual users and small and medium enterprises by informing the public about threats existing in the cyberspace and improving knowledge on the safe use of computers. Draft discussion paper on National Cyber Security Policy of Govt. of India has given importance to comprehensive national awareness program (MIT, 2011). USA has started National Initiative for Cyber Security Education (NICE, 2010) for establishing an

operational, sustainable and continually improving cyber security education program for multiple segments of the nation on correct application of sound cyber practices. Indian cyber security strategy also identifies major actions and initiatives for user awareness, education and training.

2.1.7. Motivation, reward and punishment

Information security measures are effective only when employees adhere to the security policies and practices. It has been seen that even when the security policies are well defined and made clear, the compliance level of employees towards these policies and practices may not be satisfactory due to various reasons. Reward and punishment are the two most powerful motivators (Parker, 2002). Role of penalties is a motivating factor responsible for information security behaviour (Herath & Rao, 2009). ISO/IEC 27001 also advocates that there should be a formal disciplinary process for employees who have committed a security breach.

2.1.8. Physical and environmental security of organization

Physical and environmental security refers to the methods and controls used to protect an organization proactively from natural or manmade threats to physical facilities and buildings where IT equipments are installed (Homeland Security, 2008). These measures are necessary to protect the facility housing system resources(NIST, 1996).The computer resources need physical protection from unauthorized access, theft, sabotage, various manmade or natural disasters such as fire, flood, earthquake, environmental hazards etc. (ISO/IEC 27001, 2005; ISO/IEC 27002:2005). Such protection would include fences, walls, security guards, cameras, intrusion detection systems and alarms, protection from electromagnetic radiations, uninterrupted power supply, proper temperature and humidity controls (Trček, 2006).

2.1.9. Information security audit, testing and certification

Information Security (IS) audit and testing by trained professionals are important to ascertain the actual level and effectiveness of information security in an organization. The goal of an IS audit is to have an independent party to determine the current level of security throughout the organization and point out any existing security gaps and deficiencies (Germany, 2008). Security testing through various methods and technical tools is an important way of finding out vulnerabilities in an information system. ISO/IEC 15408-1: 2009 lays down the general concepts and principles of IT security evaluation. Certification for information security such as ISO/IEC 27001: 2005 ensures that information security issues in an organization are addressed in a systematic manner. This is a seal of approval that the organization meets certain minimum standards of information security.

2.1.10. Compliance to legal and regulatory provisions

Legal provisions in a number of countries now make it mandatory for organizations, dealing in sensitive personal information, to take reasonable precautions and measures to ensure information security. Privacy and security rules of Health Insurance Portability and Accountability Act, Health Information Technology for Economic and Clinical Health Act, Gramm–Leach–Bliley Act, Federal Information Security Management Act of USA are examples of such laws. Section 43A of Indian Information Technology (Amendment) Act 2008 has provision for compensation if a body corporate fails to take reasonable security measures for protecting personal sensitive information. Compliance to such mandatory

provisions is a critical driver for organizations to take adequate security measures (Mcaffee, 2011; Symantec, 2011). Apart from the mandatory laws, many organizations follow various security standards such as ISO 27001:2005, Payment Card Industry Data Security Standard (PCI-DSS), Control Objectives for Information and related Technology (COBIT) etc. These laws and standards are also a driving force for organizations to take information security measures. Berends (2007) has shown that regulatory laws are a big driver for information security.

2.1.11. Incident management, business continuity planning (BCP) and disaster recovery (DR)

There is no such thing as absolute security. Even after best efforts an information security incident can occur due to various reasons. Organizations need to understand and mitigate risk. Risk management is the process that allows IT managers to balance the operational and financial costs of protective measures of IT systems and data that support their organizations' missions (NIST, 2002). In this process alternative strategies for dealing with risk are weighed and appropriate decisions are made (Hoo, 2000). Organizations deal with risk by identifying, analyzing and evaluating it (ISO/IEC 31000: 2009) using various techniques (ISO/IEC 31010:2009). ISO/IEC 27005 (2011) describes risk management in specific context of information. Mitigation of risk is one of the ways of dealing with risk (NIST, 2011). ISO/IEC JTC1/SC27 committee is developing another standard ISO/IEC 27014 for Governance of information security with a chapter on risk management.

Handling a disaster and recovering from an information security breach incident are integral parts of information security management. Organizations must establish, maintain and effectively implement these plans for emergency response, backup operations and post-disaster recovery for organizational information systems. This will ensure the availability of critical information resources and continuity of operations in emergency situations (NIST, 2006 B). ISO 27001:2009 advocates an effective information security incident management process along with an effective business continuity management approach. ISO/IEC 27002 (2005) also highlights the business continuity aspects of information security management. ISO/IEC 24762:2008 gives detailed guidelines for disaster recovery (DR) services for information and communications technology systems. ISO/IEC 27035:2011 lays out a structured and planned approach to information security incident management.

2.1.12. Developing information security culture in the organization

Parker (2002) has found that each organization, especially a business organization, has a unique culture that has a profound affect on how the organization deals with information security. This culture is determined by many factors including loss experience; interests and experience of top managers; acceptable ethics in the external and internal environments; the role that audit, industrial security, legal staff, and human resources play; the state of the economy, profitability, productivity, growth, and resources available; the scale of security efforts that are possible; and the many policies and practices that are or are not enforced. Dhillon (1997) has defined Information Security Culture (ISC) as the totality of patterns of behaviour in an organization that contribute to the protection of information of all kinds. Ramachandran, Srinivasan, & Tim (2008) have defined ISC as the employees' security related beliefs and values which manifest in employee's actions and behaviour in protecting the information of an organization. The aim of information security obedience by employees is to create a corporate culture where de-facto behaviour of users is what is acceptable (Thomson and Solms, 2005). Creating an

information security culture in an organization has also been emphasized by Johnson and Goetz (2007), Lim *et al*. (2009), Wright (2008) and Schjolberg and Ghernaouti-Hélie (2011).

3. Interpretive Structural Modelling (ISM)

A complex issue can be seen and analyzed from various angles and perspectives by various stakeholders. Visualization of the structure of a complex system helps in understanding it better. Interpretive Structural Modelling (ISM) is a technique to identify and depict relationship among various components of an issue or a problem. It is an effective tool to deal with complex issues. First proposed by J. Warfield in 1973, the methodology has been described by various authors such as Sage (1976), Warfield (1990), Moore (1994) etc. ISM provides a base for determining the hierarchy of the sub-elements of a program plan or any set of elements under consideration. ISM transforms unclear and poorly articulated mental models of systems into visible and well-defined models useful for many purposes. (Sushil *et al*., 2006). ISM has its foundations in mathematics as is evident in various references (Harary *et al.*, 1965). In this methodology an expert group identifies various elements which are relevant to the problem or the issue. This could be done through a survey or an Idea Engineering workshop. Idea Engineering is usually done with the help of structured questionnaire which involves the following steps:

1. Expression of a primitive statement of need.
2. Filling of structured questionnaire about objectives, obstacles and options etc. by the participants of idea engineering workshop.
3. Grouping of the ideas generated by the participants and removing the duplicates and ambiguities.
4. Prioritization and selection of most important ideas using voting (simple or weighted).

Unlike open brainstorming sessions, idea engineering is a silent exercise and open discussion is preferred before voting only. This method ensures generation of maximum ideas. After deciding the various elements of a problem or issue, it is important to identify the contextual relationships between the elements.

For ISM, pair-wise binary comparison of sub-elements is used to develop a directed graph of relationships. From this pair-wise comparison a Structural Self-Interaction Matrix (SSIM) is developed indicating pair-wise relationships. From the SSIM a reachability matrix is formed and transitivity is checked. Transitivity of contextual relationship essentially is a basic assumption in ISM which states that if element A is related to B and B is related to C, then A is necessarily related to C. If transitivity rule is violated any where, then SSIM is reviewed and modified by specific advice from expert group. This process is repeated till entire SSIM follows the transitivity rule. The final reachability matrix is partitioned into different levels. Based on the relationships seen in the reachability matrix, a directed graph (digraph) is drawn by indicating the serial number of the sub-elements and the direction of relationship with the help of an arrow. Transitivity relationships in the digraph are removed by examining it interactively. ISM is derived from the digraph by indicating the sub-elements in rectangular boxes with indicative relationships being shown by arrows. At the last, any conceptual inconsistency in the ISM Model is removed by reviewing and modifying it.

4. Variables affecting organizational information security environment

The idea engineering exercise with a group of experts and professionals working in the field of information security management in various organizations, government and academia was carried out.

Four participants were from industry, two were from academia and two were from government. The participants had more than 10 years of experience in the subject of information security management. The data was captured in a workshop mode in which sufficient time was invested in arriving at consensus for different scores in this model. The exercise generated following parameters: Top management commitment, identification & classification of information assets, technological tool/solutions, information security organizational structure and resources, information security policy and procedures, awareness, education & training, motivation, reward & punishment, physical & environmental security audit, testing and certification, compliance to legal/regulatory provisions, incident management, business continuity planning and disaster recovery and finally developing information security culture in an organization.

5. Structural Self-Interaction Matrix (SSIM)

Contextual relationship means that one variable helps to achieve another variable i.e. one "leads to" another. Based on this, contextual relationship and the associated direction between any two parameters (i and j) are identified. Four symbols are used to denote the direction of relationship between the parameters (i and j) as given below:

(1) V: parameter i will help to achieve parameter j.
(2) A: parameter i will be achieved by parameter j.
(3) X: parameters i and j will help achieve each other.
(4) O: parameters i and j are unrelated.

Based on the opinion of experts, Table 1 is developed.

Table 1: Structural Self-Interaction Matrix (SSIM)

	Parameters	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**
1	Top Management commitment	V	V	V	V	V	V	X	V	V	V	V
2	Identification & classification of information assets	O	V	V	V	O	O	A	O	O	O	
3	Technological tool/solutions	V	V	V	V	O	O	O	O	A		
4	Information security organizational structure and resources	V	V	V	V	V	V	X	V			
5	Information security policy & procedures	V	V	V	V	V	V	V				
6	Awareness, education & training	V	V	V	V	V	V					
7	Motivation, reward & punishment	V	V	O	O	V						
8	Physical & environmental security	O	V	V	V							
9	Audit, testing & certification	O	V	V								

10	Compliance to legal/regulatory provisions	V	A									
11	Incident management, Business continuity planning & Disaster recovery	O										
12	Developing information security culture in organization											

Reachability Matrix

The SSIM is converted into a binary matrix called the initial reachability matrix by substituting V, A, X, O by 1 and 0 in SSIM as per the following rules:

- If the (i, j) entry in the SSIM is V, then the (i, j) entry in the reachability matrix becomes 1 and the (j, i) entry becomes 0.
- If the (i, j) entry in the SSIM is A, then the (i, j) entry in the reachability matrix becomes 0 and the (j, i) entry becomes 1.
- If the (i, j) entry in the SSIM is X, then the (i, j) entry in the reachability matrix becomes 1 and the (j, i) entry also becomes 1.
- If the (i, j) entry in the SSIM is O, then the (i, j) entry in the reachability matrix becomes 0 and the (j, i) entry also becomes 0.

Table 2: Initial Reachability Matrix

Parameter	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	1	0	0	0	0	0	0	1	1	1	0
3	0	0	1	0	0	0	0	0	1	1	1	1
4	0	0	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	1	1	1	1	1	1	1	1
6	1	1	0	1	0	1	1	1	1	1	1	1
7	0	0	0	0	0	0	1	1	0	0	1	1

8	0	0	0	0	0	0	0	1	1	1	1	0
9	0	0	0	0	0	0	0	0	1	1	1	0
10	0	0	0	0	0	0	0	0	0	1	0	1
11	0	0	0	0	0	0	0	0	0	1	1	0
12	0	0	0	0	0	0	0	0	0	0	0	1

Removing transitivity

The initial reachability matrix in Table 2 is checked for the transitivity rule. This rule means that in a relationship between three parameters A, B and C, if A leads to B and B leads to C then it necessarily means that A must also lead to C. Wherever the transitivity rule is violated , SSIM is reviewed and suitably modified by giving expert feedback. Driver power and dependence of each parameter is obtained by summing up the rows and columns respectively. The result is revised reachability matrix as shown in Table 4.

Table 3: Revised Reachability Matrix

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Driver power**
1	1	1	1	1	1	1	1	1	1	1	1	1	**12**
2	0	1	0	0	0	0	0	0	1	1	1	1	**5**
3	0	0	1	0	0	0	0	0	1	1	1	1	**5**
4	1	1	1	1	1	1	1	1	1	1	1	1	**12**
5	1	1	1	1	1	1	1	1	1	1	1	1	**12**
6	1	1	1	1	1	1	1	1	1	1	1	1	**12**
7	0	0	0	0	0	0	1	1	1	1	1	1	**6**
8	0	0	0	0	0	0	0	1	1	1	1	1	**5**
9	0	0	0	0	0	0	0	0	1	1	1	1	**4**
10	0	0	0	0	0	0	0	0	0	1	0	1	**2**

11	0	0	0	0	0	0	0	0	0	1	1	1	**3**
12	0	0	0	0	0	0	0	0	0	0	0	1	**1**
Dependence	**4**	**5**	**5**	**4**	**4**	**4**	**5**	**6**	**9**	**11**	**10**	**12**	

Level partitions

Final reachability matrix leads to the reachability and antecedent set for each parameter (Warfield, 1974). The reachability set consists of the parameter itself and the other parameter that it may impact, whereas the antecedent set consists of the element itself and the other parameter that may impact it. Intersection of these sets is derived for all the parameters. The parameters for whom the reachability and the intersection sets are the same occupy the top level in the ISM hierarchy. The top-level parameter in the hierarchy would not lead to any other parameter above its own level. After identifying the top-level parameter, it is separated out from the other parameters. Same process is repeated for finding the top-level parameters in the next level. The process is continued until the level of each parameter is obtained (Tables 4 to 10). These levels help in building the digraph, hierarchy and the final model.

Table 4: Iteration 1

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,2,3,4,5,6,7,8,9,10,11,12	1,4,5,6	1,4,5,6	
2	2,9,10,11,12	1,2,4,5,6	2	
3	3,9,10,11,12	1,3,4,5,6	3	
4	1,2,3,4,5,6,7,8,9,10,11,12	1,4,5,6	1,4,5,6	
5	1,2,3,4,5,6,7,8,9,10,11,12	1,4,5,6.	1,4,5,6	
6	1,2,3,4,5,6,7,8,9,10,11,12	1,4,5,6	1,4,5,6	
7	7,8,9,10,11,12	1,4,5,6,7	7	
8	8,9,10,11,12	1,4,5,6,7,8	8	
9	9,10,11,12	1,2,3,4,5,6,7,8,9	9	
10	10,12	1,2,3,4,5,6,7,8,9,10,11	10	

11	10,11,12	1,2,3,4,5,6,7,8,9,11	11	
12	12	1,2,3,4,5,6,7,8,9,10,11,12	12	I

As evident from the above table 4, parameter 12 does not lead to any other variable and hence it is marked at level I and is subsequently removed in the next iteration leading to table 5 below.

Table 5: Iteration 2

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,2,3,4,5,6,7,8,9,10,11	1,4,5,6	1,4,5,6	
2	2,9,10,11	1,2,4,5,6	2	
3	3,9,10,11	1,3,4,5,6	3	
4	1,2,3,4,5,6,7,8,9,10,11	1,4,5,6	1,4,5,6	
5	1,2,3,4,5,6,7,8,9,10,11	1,4,5,6.	1,4,5,6	
6	1,2,3,4,5,6,7,8,9,10,11	1,4,5,6	1,4,5,6	
7	7,8,9,10,11	1,4,5,6,7	7	
8	8,9,10,11	1,4,5,6,7,8	8	
9	9,10,11	1,2,3,4,5,6,7,8,9	9	
10	10	1,2,3,4,5,6,7,8,9,10,11	10	II
11	10,11	1,2,3,4,5,6,7,8,9,11	11	

As evident from the above table 5, parameter 10 does not lead to any other variable and hence it is marked at level II and is subsequently removed in the next iteration leading to table 6 below.

Table 6: Iteration 3

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,2,3,4,5,6,7,8,9,11	1,4,5,6	1,4,5,6	
2	2,9,11	1,2,4,5,6	2	
3	3,9,11	1,3,4,5,6	3	
4	1,2,3,4,5,6,7,8,9,11	1,4,5,6	1,4,5,6	
5	1,2,3,4,5,6,7,8,9,11	1,4,5,6.	1,4,5,6	
6	1,2,3,4,5,6,7,8,9,11	1,4,5,6	1,4,5,6	
7	7,8,9,11	1,4,5,6,7	7	
8	8,9,11	1,4,5,6,7,8	8	
9	9,11	1,2,3,4,5,6,7,8,9	9	
11	11	1,2,3,4,5,6,7,8,9,11	11	III

As evident from the above table 6, parameter 11 does not lead to any other variable and hence it is marked at level III and is subsequently removed in the next iteration leading to table 7 below.

Table 7: Iteration 4

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,2,3,4,5,6,7,8,9	1,4,5,6	1,4,5,6	
2	2,9	1,2,4,5,6	2	
3	3,9	1,3,4,5,6	3	
4	1,2,3,4,5,6,7,8,9	1,4,5,6	1,4,5,6	
5	1,2,3,4,5,6,7,8,9	1,4,5,6.	1,4,5,6	
6	1,2,3,4,5,6,7,8,9	1,4,5,6	1,4,5,6	

7	7,8,9	1,4,5,6,7	7	
8	8,9	1,4,5,6,7,8	8	
9	9	1,2,3,4,5,6,7,8,9	9	IV

As evident from the above table 7, parameter 9 does not lead to any other variable and hence it is marked at level III and is subsequently removed in the next iteration leading to table 8 below.

Table 8: Iteration 5

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,2,3,4,5,6,7,8	1,4,5,6	1,4,5,6	
2	2	1,2,4,5,6	2	V
3	3	1,3,4,5,6	3	V
4	1,2,3,4,5,6,7,8	1,4,5,6	1,4,5,6	
5	1,2,3,4,5,6,7,8	1,4,5,6.	1,4,5,6	
6	1,2,3,4,5,6,7,8	1,4,5,6	1,4,5,6	
7	7,8	1,4,5,6,7	7	
8	8	1,4,5,6,7,8	8	V

As evident from the above table 8, parameters 2,3 and 8 do not lead to any other variable and hence they are marked at level IV and are subsequently removed in the next iteration leading to table 9 below.

Table 9: Iteration 6

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,4,5,6,7	1,4,5,6	1,4,5,6	
4	1,4,5,6,7	1,4,5,6	1,4,5,6	

5	1,4,5,6	1,4,5,6.	1,4,5,6	
6	1,4,5,6,7	1,4,5,6	1,4,5,6	
7	7	1,4,5,6,7	7	VI

As evident from the above table 9, parameter 7 does not lead to any other variable and hence it is marked at level VI and is subsequently removed in the next iteration leading to table 10 below.

Table 10: Iteration 7

Variable	Reachability set	Antecedent set	Iteration set	Level
1	1,4,5,6	1,4,5,6	1,4,5,6	VII
4	1,4,5,6	1,4,5,6	1,4,5,6	VII
5	1,4,5,6	1,4,5,6.	1,4,5,6	VII
6	1,4,5,6	1,4,5,6	1,4,5,6	VII

ISM model

Relationship of various top-level parameters from level I to VII is indicated. If there is a relationship between the parameter i and j, this is shown by an arrow which points from i to j. This graph is called a directed graph, or digraph (Figure 1). After removing the transitivities the digraph is finally converted into the ISM-based model (Figure 2).

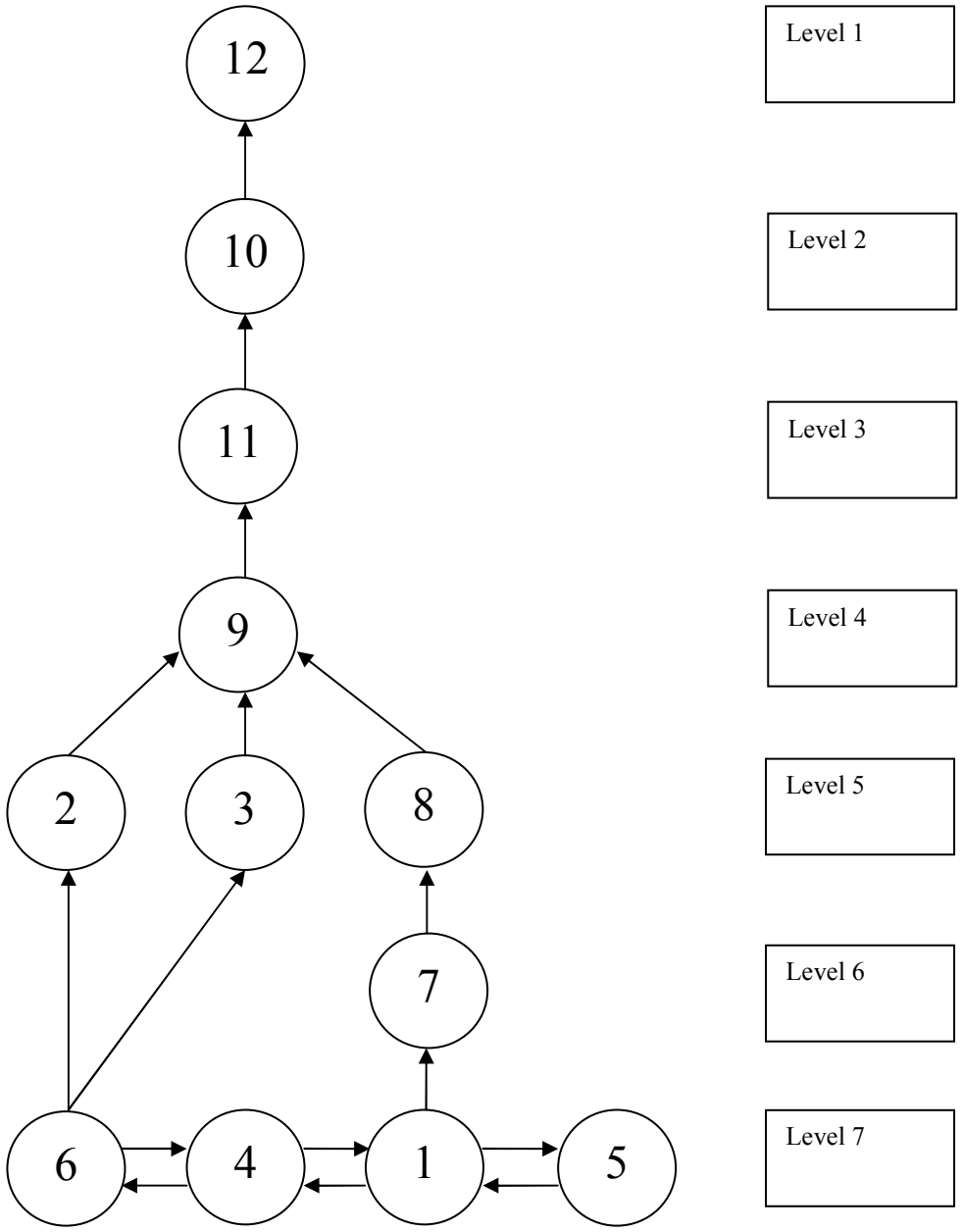


Fig. 1: Diagraph for parameters

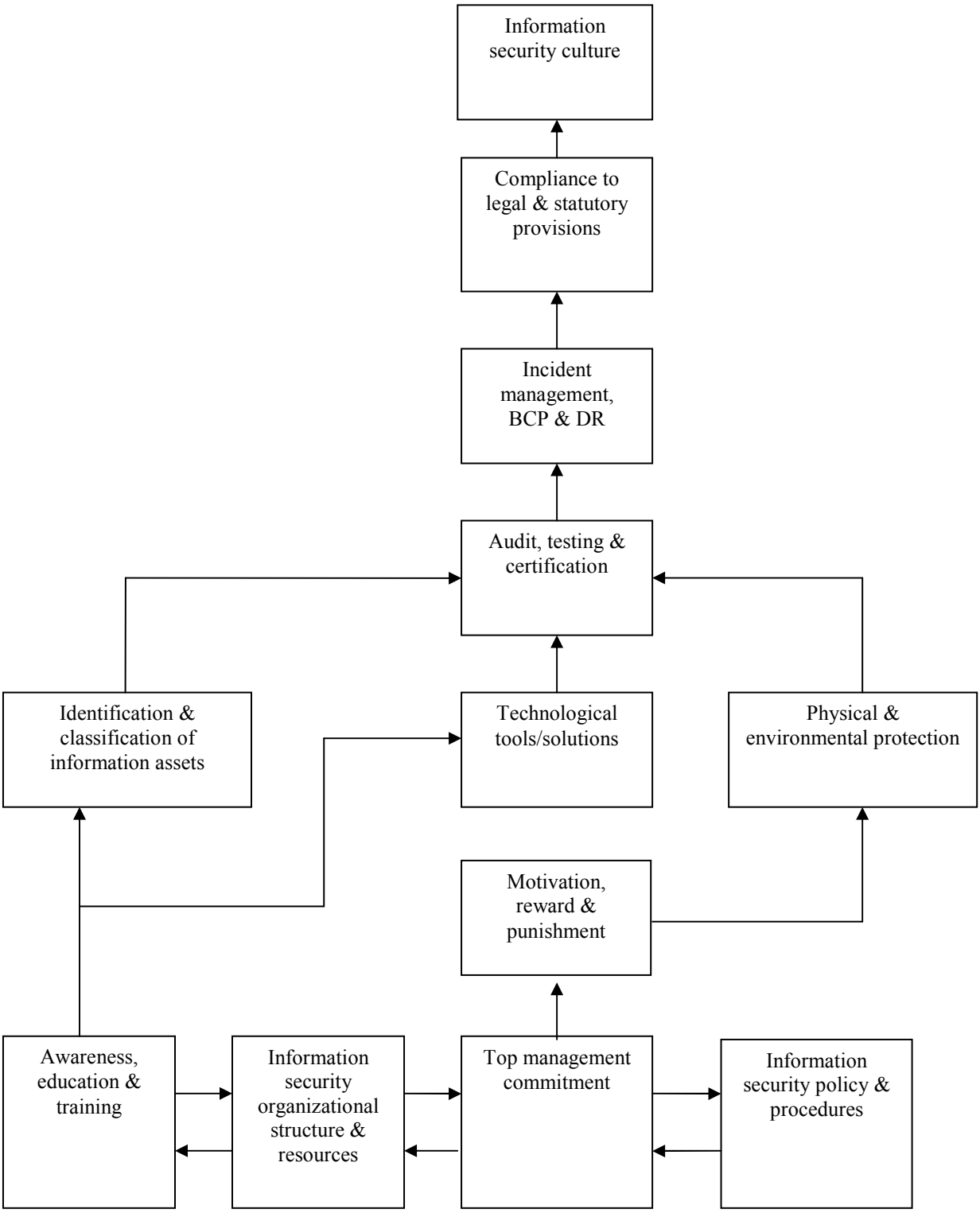


Fig. 2: Interpretive structural model

6. MICMAC Analysis

Cross-Impact Matrix Multiplication Applied to Classification (MICMAC) was developed by Godet (1993).The MICMAC principle is based on multiplication properties of matrices (Sharma *et al*. 1995). The objective of MICMAC analysis is to classify the variables according to their driving power and dependence. The variables are plotted on a graph with dependence on X axis and driving power on Y axis. The variable can be classified in four categories. The first category is of the variables with low dependence and low driving power. These variables are called autonomous or excluded variables. The second category variables are those which have high dependence but low driving power. These are called dependent variables. The third category is of those which have high dependence as well as driving power. These are called relay variables as any effect on them will affect higher level variables but they themselves are dependent on lower level variables. The fourth category variables have low dependence but high driving power and therefore are called driving variables. These variables are the most important variables as they strongly influence the other variables. They are also considered as the entry level variables in the system. Based on the driver power and dependence of various parameters in Table 3, a driving power- dependence diagram is constructed as shown in Figure 3

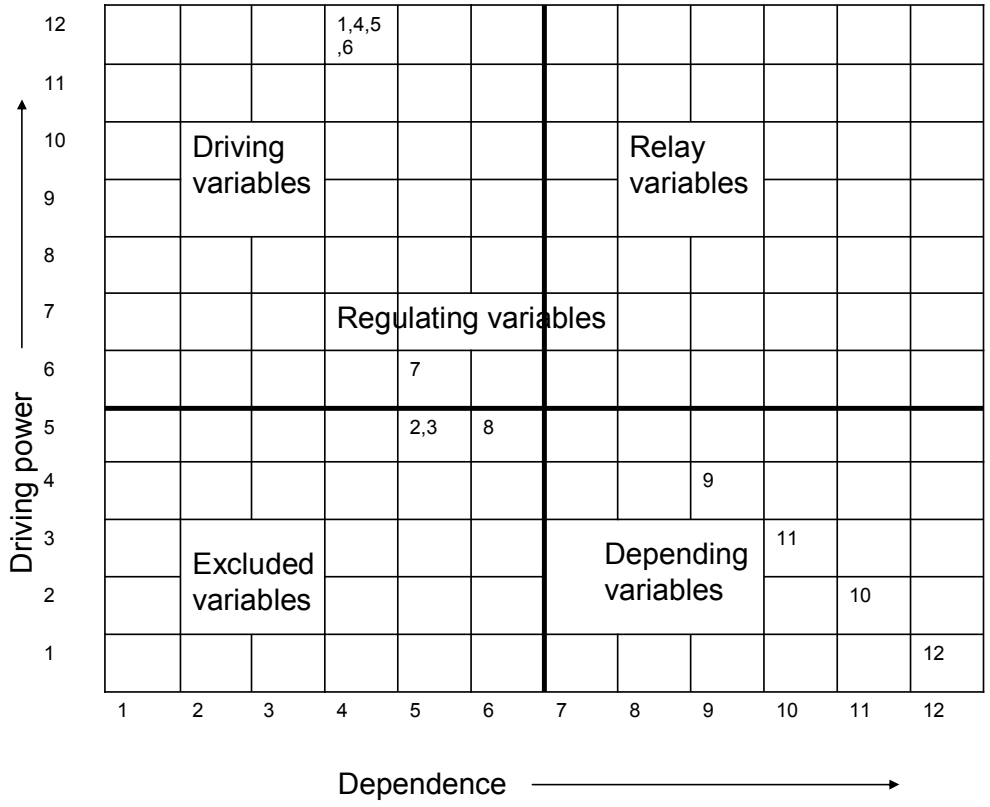


Fig. 3: Driving power and dependence diagram

7. Discussions and conclusions

This study has attempted to study the inter-relationships among various information security practices and the underlying structure as was suggested by Ma and Pearson (2005). Several interesting and useful findings about the implementation of information security in an organization can be drawn from the application of interpretive structural modelling approach. The driving power-dependence diagram as shown in Figure 3 has helped to classify various factors affecting information security environment in an organization into four distinct clusters. These clusters can be analyzed based on their driving power and dependence. The first cluster is of those variables which have low driving power and low dependence. In this analysis even though variable such as 'identification & classification of information assets', 'technological tool/solutions', 'physical & environmental protection' and 'motivation, reward & punishment' have relatively low driving power and dependence but they almost fall near the middle of the diagram. These are regulating variables situated mostly at the system's center of gravity. Being quite above the diagonal they also become secondary levers and are more influent than dependent (Arcade *et al.*, 2003). The second cluster located in the south-east frame of the diagram consists of variables which have low driving power but high dependence. Being result variables they are influenced by driving variables situated in the north-west frame of the diagram. In present case 'audit, testing and certification', 'compliance to legal & statutory provisions, 'incident management, BCP & DR' and 'development of information security culture' are the variables which fall in this category. There are no variable which have high driving power and high dependence. The forth cluster is of those variables which have high driving power and low dependence. These factors are key success parameters or enablers for achieving the end results. These are 'top management commitment', 'information security organizational structure', 'information security policy and procedures' and 'awareness, education and training'. As revealed by the diagraph, these factors are also mutually affecting each other and are closely linked indicating that equal emphasis is required on each of them to achieve information security in an organization. Since no autonomous variables have been found in the above analysis, it is proved that all variable influence information security. These conclusions are highly useful for strategic decision makers in an organization. The graphic model is of importance in understanding how various factor affecting information security in an organization interact and influence each other. Kumar (2009), while discussing the case study of Bharat Heavy Electrical Ltd., a government of India undertaking, has mentioned that corporate information technology was the driver for corporation wise implementation. ISM model also reveals that technological tools and solutions and information security policy and procedures have a very high driving power during successful implementation of information security. This was also established through MICMAC analysis (figure 3). Kumar (2009) also reports that employee awareness and education has a major driving power which corroborates our findings. Alkaff (2009) points out that top management support has led to Abhu Dhabi Gas Industry Ltd. (GASCO) to achieve implementation of ISO/IEC 27001. This also corroborates our results of ISM (figure 2). ISM model has revealed that operational variables related to implementation such as identification and classification of information assets, use of technological tools/solutions and physical and environmental protection are driven through top management support.

8. Limitations and scope for future work

In the present study, 12 parameters were identified based on literature study and expert help. It is possible to identify some more parameters for ISM development. Help of experts was also taken to identify the contextual relationship among the variables for ISM model. This may introduce some element of bias.

Although a relationship model using ISM has been developed but it has not been validated statistically. For future research, it is suggested that structural equation modelling (SEM) technique may be used to corroborate the findings of ISM. Some of the variables have been clubbed together being a part of a subset due to their similar nature but it is possible to treat them as independent variables. Future researches may establish their interrelationships also.

9. References

[1] Al-Awadi, 2009, "A study of Employees' Attitudes Towards Organizational Information Security Policies in the UK and Oman"

[2] Alkaff, Adel Salem, 2009, "Case studies show value of ISO/IEC 27001 conformity", *ISO Management Systems* January-February 2009 pp. 33-38.

[3] Arcade, Jaques;Godet, Michel;Meunier,F.; Roubelat, F., 2003, "Structural Analysis with MICMAC Method & Actors' Strategy with MACTOR Method", available at http://www.agriperi.ir/AKHBAR/cd1/FORESIGHT%20METHODOLOGY%20&%20FORECASTING/FORESIGHT%20METHODOLOGY/related%20articles/books/Future%20Research%20Methodology/4-stranal.pdf

[4] Australia, 2009, "Cyber Security Strategy", available at http://www.ag.gov.au/www/agd/rwpattach.nsf/VAP/%284CA02151F94FFB778ADAEC2E6EA8653D%29~AG+Cyber+Security+Strategy+-+for+website.pdf/$file/AG+Cyber+Security+Strategy+-+for+website.pdf

[5] Avolio, F.M. , 2000, " Best Practices in Network Security", Network Computing, 60(20), March.

[6] Belasco, Kent and Wan, Siaw-Peng, 2006, " Online Retail Banking: security Concerns, Breaches and Controls", Handbook of Information Security Vol 3 page 43: New Jersey, John Wiley and sons. Inc.

[7] Berends A. J. , 2007, "Dealing with information loss" Masters Thesis available at http://www.tbm.tudelft.nl/fileadmin/Faculteit/TBM/Over_de_Faculteit/Afdelingen/Afdeling_Infrastructure_Systems_and_Services/Sectie_Informatie_en_Communicatie_Technologie/medewerkers/jan_van_den_berg/news/doc/arjen-berends.pdf

[8] Björck, Fredrik, 2001, "Implementing Information Security Management Systems - An Empirical Study of Critical Success Factors" available at http://people.dsv.su.se/~bjorck/files/success-factors.pdf

[9] CERT, 2009, "CERT Research annual report 2009", available at http://www.cert.org/research/2009research-report.pdf

[10] CSO, 2010, "Cyber Security watch Survey Results", available at http://www.csoonline.com/documents/pdfs/2010CyberSecurityResults.pdf

[11] Dhillon, G., 1997, *Managing Information System Security*. Houndmills, Basingstoke, Hampshire: Macmillan Press Ltd.

[12] Draft ISO Guide 83, "High level structure and identical text for management system standards and common core management system terms and definitions".

[13] Englesman, Wilco, 2007, "Information Assets and their Value", available at http://referaat.cs.utwente.nl/new/paper.php?paperID=186

[14] ENISA, 2009, "The growing requirement for information security awareness" available at http://www.enisa.europa.eu/act/ar/deliverables/2009/ar-book-09/at_download/fullReport

[15] Estonia , 2008, " Cyber Security Strategy", available at http://www.mod.gov.ee/files/kmin/img/files/Kuberjulgeoleku_strateegia_2008-2013_ENG.pdf

[16] Gallaher Michael P., Link, Albert N. and Rowe, Brent R.,2008, "Cyber security: Economic Strategies and Policy Alternatives" Pages 41-55 :UK, Edward Elgar Publishing Ltd.

[17] Germany, 2008, "Information security audit (IS audit) - A guideline for IS audits based on IT-Grundschutz" available at https://www.bsi.bund.de/cae/servlet/contentblob/471376/publicationFile/27987/guideline-isrevision_pdf.pdf

[18] Godet, M., (1993), "From anticipation to action: a handbook of strategic prospective", UNESCO Publishing.

[19] Harary F., Norman R.Z. and Cartwright D., 1965, "Structural Models: An Introduction to the Directed Graphs": New York, Wiley.

[20] Harris, Shon, "All in one CISSP Exam Guide", 5th edition: Noida, Tata MacGraw Hill Education.

[21] Herath, Tejaswini and Rao, H.R., 2009, "Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness", Decision Support Systems Volume 47, Issue 2, May 2009, Pages 154-165.

[22] Homeland Security, 2008, "Information Technology (IT) Security Essential Body of Knowledge (EBK): A Competency and Functional Framework for IT Security Workforce Development" available at http://www.us-cert.gov/ITSecurityEBK/EBK2008.pdf

[23] Hong, Kwo-Shing, Chi,Yen-Ping, Chao,Louis R.,Tang Jih-Hsing, "An empirical study of information security policy on information security elevation in Taiwan", *Information Management & Computer Security* Vol. 14 No. 2, 2006 pp.104-115.

[24] Hoo, Kevin J. Soo, 2000, "How Much is Enough? A Risk-Management Approach to Computer Security", , Stanford, Center for International Security and Cooperation (CISAC)

[25] ISO/IEC 27001: 2005, " Information technology-Security techniques-Information security management systems-Requirements"

[26] ISO/IEC 27001: 2009, "Information technology— Security techniques — Information security management systems — Overview and vocabulary"

[27] ISO/IEC 27002:2005, "Information technology — Security techniques — Code of practice for information security management"

[28] ISO/IEC 27005:2011, "Information technology— Security techniques — Information security risk management."

[29] ISO/IEC 31000:2009, "Risk management — Principles and guidelines"

[30] ISO/IEC 31010:2009, "Risk management — Risk Assessment techniques"

[31] ISO/IEC 27035:2011, "Information technology— Security techniques — Information security incident management"

[32] Johnson, M. Eric, Goetz, Eric, 2007, "Embedding Information Security into the Organization", IEEE security and privacy May-June 2007 page 16-24.

[33] Kankanhalli Atreyi, Hock-hai Teo, Bernard C.Y.Tan, and Kwok-kee Wei, 2003, " An integrative study of information systems security effectiveness", International Journal of Information Management, 2003 pages 139-154.

[34] Kajava, J., Anttila, J.,Varonen, R.,Savola, R., Roning,J., " Senior Executives Commitment to Information Security-from Motivation to Responsibility" proceedings of International Conference, CIS 2006, Guangzhou, China, November 3-6, 2006.

[35] Knapp,Kenneth J., Marshall, Thomas E., Rainer, R.K., Morrow, Dorsey W., 'Information Systems Security, 2006 pages 51-58.

[36] Kumar, Arvind, ISO, "Case studies show value of ISO/IEC 27001 conformity", *ISO Management Systems* January-February 2009 pp. 33-38.

[37] Lim, Joo S, Chang, Maynard, Sean, Ahmad, Atif, 2009, "Exploring the Relationship between Organizational Culture and Information Security Culture", Proceedings of the 7th Australian Information Security Management Conference.

[38] Ma, Q. and Pearson, J.M., 2005, "ISO 17799: Best Practices in Information Security Management?", Communications of the AIS,15,577-591.

[39] Mcafee, 2011, "Risk and Compliance Outlook", available at http://www.mcafee.com/us/resources/reports/rp-risk-compliance-outlook-2011.pdf

[40] MIT, 2011, "Discussion draft on National Cyber Security Policy", Department of Information Technology,Ministry of Communications and Information Technology Government of India, available at http://www.mit.gov.in/sites/upload_files/dit/files/ncsp_060411.pdf

[41] Moore, C. M., 1994, *Group Techniques for Idea Building*: Beverly Hills Sage Publications.

[42] NICE, 2010, "National Initiative for Cyber Security Education", available at http://www.whitehouse.gov/sites/default/files/rss_viewer/cybersecurity_niceeducation.pdf

[43] NIST, 1996, "Generally Accepted Principles and Practices for Securing Information Technology Systems", NIST 800-14 available at http://csrc.nist.gov/publications/nistpubs/800-14/800-14.pdf

[44] NIST, 2002, "Risk Management Guide for Information Technology Systems", NIST Special Publication 800-30, available at http://csrc.nist.gov/publications/nistpubs/800-30/sp800-30.pdf

[45] NIST, 2006 A, "Information Security handbook: A Guide for Managers", NIST Special Publication 800-100, available at http://csrc.nist.gov/publications/nistpubs/800-100/SP800-100-Mar07-2007.pdf

[46] NIST, 2006 B, "Guide for Developing Performance Metrics for Information Security", NIST Special Publication 800-80, available at http://permanent.access.gpo.gov/lps72067/draft-sp800-80-ipd.pdf

[47] NIST, 2008, "Performance Measurement Guide to Information Security", NIST Special Publication 800-55 rev. 1, available at http://csrc.nist.gov/publications/nistpubs/800-55-Rev1/SP800-55-rev1.pdf

[48] NIST, 2009, "Recommended Security Controls for Federal Information Systems and Organizations", NIST Special Publication 800-53 revision 3, available at http://csrc.nist.gov/publications/nistpubs/800-53-Rev3/sp800-53-rev3-final.pdf

[49] NIST, 2011, "Managing Information Security Risk", NIST Special Publication 800-39, available at http://csrc.nist.gov/publications/nistpubs/800-39/SP800-39-final.pdf

[50] OECD, 2002, "OECD Guidelines for the Security of Information Systems and Networks: Towards a Culture of Security", available at http://www.oecd.org/dataoecd/16/22/15582260.pdf

[51] Parker, Donn, B., 2002, "Motivating the Workforce to Support Security Objectives:A Long-Term View" available at http://www.iwar.org.uk/comsec/resources/sa-tools/Motivation-for-Information-Security.pdf

[52] Ramachandran, S., Srinivasan, V. R., & Tim, G. (2008). *Information Security Cultures of Four Professions: A Comparative Study.* In Proceedings of the 41st Hawaii International Conference on System Sciences - 2008, Hawaii.

[53] Sage, A.P., 1976, *Methodology for Large Scale Systems*: New York, Mc Graw-Hill.

[54] Sami, Abu-Zineh , 2006, "Success Factors of Information Security Management: A Comparative Analysis between Jordanian and Finnish Companies", available at http://www.pafis.shh.fi/graduates/samabu04.pdf

[55] Schjolberg, Stein and Ghernaouti-Hélie, Solange, 2011," A Global Treaty on Cybersecurity and Cybercrime", 2nd Edition, available at

http://www.cybercrimelaw.net/documents/A_Global_Treaty_on_Cybersecurity_and_Cybercrime,_Second_edition_2011.pdf

[56] Sharma, H.D., Gupta, A.D. and Sushil, "The objectives of waste management in India: a future inquiry". Tech. Forecast. & Social Change, 1995, 48, 285–309.

[57] Singh, M.D. & Kant, R., 2008, "Knowledge management barriers: An interpretive structural modeling approach", International Journal of Management Science and Engineering Management, Vol. 3, No. 2, pp. 141-150.

[58] Siponen, T. 2000, "A conceptual foundation for organizational information security awareness", Information Management & Computer Security 8/1 2000, 31-41.

[59] Sushil, Saxena J. P., Prem Vrat, 2006, "Policy and Strategy Formulation: An Application of Flexible Systems Methodology", Delhi: GIFT Publishing.

[60] Symantec, 2011, "Symantec Security Check – Indian Financial Services Industry 2011", available at http://www.symantec.com/content/en/in/enterprise/collateral/other_resources/MediaPresentation_SymantecSecurityCheck_Indian%20FinancialServices_SurveyFindings_August18.pdf

[61] Thomson, Kerry-Lynn and Solms, R. von, 2005, *Computers & Security* (2005) 24, 69-75.

[62] Trček, Denis, 2006, "Managing Information Systems Security and Privacy", Berlin: Springer.

[63] Tudor J. K., 2001, "IS security Architecture, An Integrated Approach to Security in the Organization." USA: Auerbach Publications.

[64] UK, 2009, "Cyber Security Strategy of the United Kingdom :safety, security and resilience in cyber space", available at http://www.cabinetoffice.gov.uk/media/216620/css0906.pdf

[65] Vaish, Abhishek and Varma, Shirshu, 2010, "Parameter Extraction for Measurement of the Effective Information Security Management - Statistical Analysis" International Journal of Computer and Electrical Engineering, Vol. 2, No. 4, August, 2010,1793-8163

[66] Von Solms, B and R. Von Solms, 2004, "The 10 deadly sins of Information Security Management", *Computer & Security,* Vol. 23, No. 3, pp. 191-8.

[67] Von Solms, B, 2006, "Information Security-The Forth Wave", *Computers & Security* 25 (2006) 165–168

[68] Warfield, J. N., 1990, *A Science of Generic Design: Managing Complexity through System Design*, Vol. I & II, International Systems, Salinas.

[69] Warfield, J.W. (1974), "Developing interconnected matrices in structural modeling", IEEE Transactions on Systems Men and Cybernetics, Vol. 4 No. 1, pp. 51-81.

[70] Wright, Craig, 2008, "The IT Regulatory and Standards Compliance Handbook" USA: Syngress Publishing Inc.

Novel Applications of Functionally Graded Nano, Optoelectronic and Thermoelectric Materials

Isha Bharti, Nishu Gupta, and K. M. Gupta, *Member, IACSIT*

***Abstract*—Functionally graded materials are a novel class of materials having unique characteristics. They are graded property materials and are used as medical implants, for thermal protection of space vehicles, as thermoelectric converter for energy conservation etc. Due to their versatility of behaviour, they are now used as nano, optoelectronic and thermoelectric materials also. Future applications demand materials having extraordinary mechanical, electronic and thermal properties which can sustain different environment conditions and are easily available at reasonable prices. The carbon nanotubes (CNT) reinforced functionally graded composite materials (FGCM) is expected to be the new generation material having a wide range of unexplored potential applications in various technological areas such as aerospace, defence, energy, automobile, medicine, structural and chemical industry. They can be used as gas adsorbents, templates, actuators, catalyst supports, probes, chemical sensors, nanopipes, nano-reactors etc.**

This paper aims at elaborating the development of such materials by compiling the ongoing researches. In this regard, the research developments of newer materials by different investigators have been presented here. Brief details of the development of CNT reinforced Functionally Graded composite materials, FGM in optoelectronic applications, FG thermoelectric materials have been presented here. In these elaborations, it is shown by the respective investigators that these FGMs can be effectively used as MRI scanner cryogenic tubes, tools and dies, solar panels, photodetectors , high temperature joining materials and various structural elements etc.

***Index Terms*—CNT reinforced functionally graded composite materials (FGCMs), FGM in optoelectronics, FG thermoelectric materials.**

I. Introduction to Functionally Graded Materials (FGMS)

A functionally graded material (FGM) is a two-component composite characterised by a compositional gradient from one component to other. In contrast to traditional composites that are homogeneous mixtures and involve a compromise between the desirable properties of component materials; the significant proportions of FGMs contain the pure form of each component, hence the need for compromise is eliminated. Thus the properties of both components can be fully utilised. For example, the toughness of a metal can be mated with the refractoriness of a ceramic without any compromise in the toughness of the metal or the refractoriness of the ceramic.

Functionally graded materials may be characterized by the variation in their composition and structure gradually over the volume, resulting in corresponding changes in their properties. The materials can be designed for specific functions and applications.

Functionally graded materials (FGMs) are a new generation of engineered materials that are gaining interest in recent years. FGMs were initially designed as thermal barrier materials for aerospace structural applications and fusion reactors. FGMs also found applications in structural components operating under extremely high-temperature environment. As an example, FGMs based on ceramic reinforcement in metal matrix are able to withstand high-temperature environments due to better thermal resistance of ceramic constituents, while the metal constituents enhance their mechanical performance and reduce the possibility of catastrophic fracture. FGMs are the composite materials in which the content of reinforcement is gradually varied in some direction to achieve gradient in properties. Due to graded variation in the content of constituent materials, the properties of FGMs undergo appreciable and continuous change from one surface to another, thus eliminating interface problems and diminishing thermal stress concentrations. There are many areas of application for FGMs. The concept is to make a composite material by varying the microstructure from one material to another material with a specific gradient. This enables the material to have the best of both materials. If it is for thermal or corrosive resistance, or malleability and toughness, both strengths of material may be used to avoid corrosion, fatigue, fracture and stress corrosion cracking.

A. Types of FGMs

Depending upon the nature of gradient, the functionally graded materials (composites) may be grouped into following types [1].

1) Fraction gradient type (Fig. 1a)
2) Shape gradient type (Fig.1b)
3) Orientation gradient type (Fig.1c)
4) Size (of material) gradient type (Fig.1d)

II. CNT Reinforced FGM Composites [2]

CNT reinforced metal or ceramic matrix functional graded composites exhibit improvements in properties such as thermo- mechanical, lightweight, dimensional stability, barrier properties, flame retardancy heat resistance and

Manuscript received January 4, 2013; revised March 4, 2013.

Isha Bharti is with the Department of Applied Physics, Delhi Technological University, Delhi (e-mail: dtu.isha@gmail.com).

Nishu Gupta, was with the Department of Electronics and Communication Engineering, Motilal Nehru National Institute of Technology, Allahabad (e-mail: dce.nishu@gmail.com).

K. M. Gupta is with the Department of Applied Mechanics, Motilal Nehru National Institute of Technology, Allahabad (e-mail: kmgupta@mnnit.ac.in).

DOI: 10.7763/IJMMM.2013.V1.47

challenging technology in future. Compared to metals these show better higher strength-to- density ratios, higher stiffness-to-density ratios, better fatigue and wear resistance, better elevated temperature properties (higher strength-lower creep rate. In contrast to composites materials, FG-composites have new unique properties like fabricate directional mechanical properties, higher temperature capability (lower thermal expansion properties), and excellent fatigue and fracture resistance. However, it is difficult to meet stringent dimensional stability requirements during gradation and results in higher cost of the product.

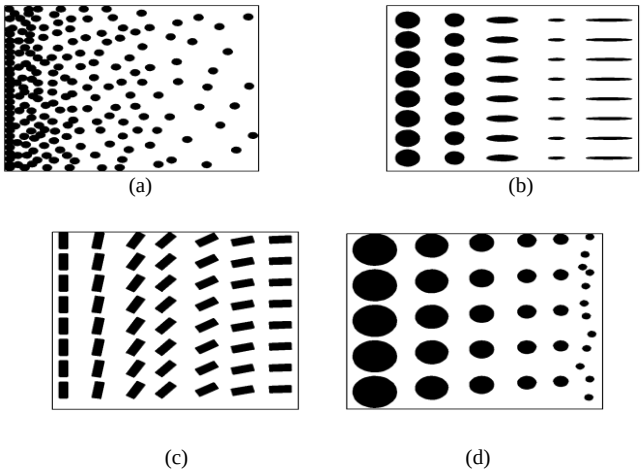


Fig. 1. Based on vividities of gradients, different types of functionally graded materials may be of (a) fraction gradient type, (b) shape gradient type,(c) orientation gradient type, and (d) size gradient type [1].

A. Applications of CNT in FGM

CNT reinforced metal matrix functional graded-composites due to their unique combination of hardness, toughness and strength are universally used in cutting tools, drills, machining of wear resistant materials, mining and geothermal drilling. CNT reinforced functional graded-composite materials have the ability to generate new features and perform new functions that are more efficient than larger structures and machines. Due to functional variation of FGM-materials, their physical/chemical properties (e.g. stability, hardness, conductivity, reactivity, optical sensitivity, melting point, etc.) can be manipulated to improve the overall properties of conventional materials. Some of the current and futuristic applications of FGM are listed as follows [2].

1) Commercial *and Industrial:* Pressure vessels, Fuel tanks, Cutting tool inserts, Laptop cases, Wind turbine blades, Firefighting air bottles, MRI scanner cryogenic tubes, Eyeglass frames, Musical instruments, Drilling motor shaft, X-ray tables, Helmets.
2) Automobiles*:* (SiC-SiC) Combustion chambers,(AI-SiC) Engine cylinder liners, (SiCw/AI-alloy) Diesel engine pistons, (E-glass/epoxy) Leaf springs, (AI-C) Drive shafts, (AI-SiC) Flywheels, Racing car brakes, (SiCp/AI-alloy) Shock absorbers.
3) Aerospace *equipment and structures:* (TiAl-SiC fibers) Rocket nozzle, Heat exchange panels, Spacecraft truss structure, Reflectors, Solar panels, Camera housing, Turbine wheels (operating above 40,000 rpm),Nose caps and leading edge of missiles and Space shuttle.
4) Sub-*marine:* (Carbon and glass fibers) Propulsion shaft, (Graphite/Epoxy) Cylindrical pressure hull, (Glass/Epoxy) Sonar domes, Composite piping system, (AI-SiC) diving cylinders.
5) Biotechnology: The functional gradient nanohydroxyapatite reinforced polyvinyl alcohol (nanoHA/PVA) gel biocomposites can be prepared through a layer-by-layer casting method combining with freeze/thaw cycles technology. NanoHA/PVA gel biocomposites have been a promising and excellent artificial articular cartilage repair material. Compared to hydrogel-based materials such as poly(vinyl alcohol) (PVA), poly(lactic acid) and chitosan, nanoHA/PVA gel biocomposites possesses superior bioactivity and mechanical properties because of the nanoHA existence in the composites. [3], [4]

B. Other Applications of CNT Reinforced FGM

Potential applications of FGM are both diverse and numerous. Some more applications of CNT in FGM having recent applications are the following [2]:

1) CNT reinforced functionally graded piezoelectric actuators.
2) As furnace liners and thermal shielding elements in microelectronics.
3) CNT reinforced functionally graded tools and dies for better thermal management, better wear resistance, reduce scrap, and improved process productivity.
4) CNT reinforced functionally graded polyester-calcium phosphate materials for bone replacement.

III. FGM in Optoelectronic Devices [5]

Functionally graded materials (FGMs) are materials in which some particular physical properties are changed with dimensions. Properties of such materials can be described by material function f(*x*). In homogenous materials this function is constant, like in Fig. 2a. In case of a junction of two different materials, function f(x) has a stair-shape (Fig. 2b).

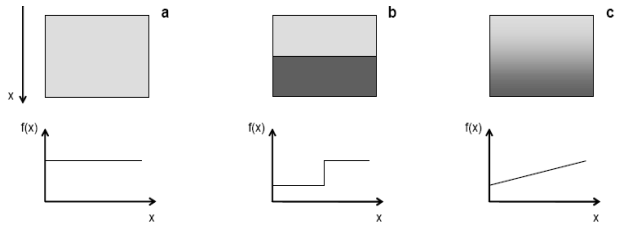


Fig. 2. Schematic representation of materials function in different structures; homogenous material (a), junction (b), FGM (c)

In FGM, this material function should be continuous or quasi-continuous. It means that the particular properties change continuously or quasi continuously along one direction, like it is shown in Fig. 2**c**. In many cases, FGM can be presented as a composition of several connected thin layers.

Depending on the number of directions the proprieties changed, we can discriminate as 1- dimensional, 2-dimensional or 3-dimensional FGM. It can be mathematically described for 3-D FGM as

$$\frac{dF}{dx} \neq 0, \quad \frac{dF}{dy} \neq 0, \quad \frac{dF}{dz} \neq 0$$

where $F(x, y, z)$ is the material function.

A. Possible Applications of FGM in Optoelectronics

Now-a-days the graded materials are widely used for anti-reflective layers, fibers, GRIN lenses and other passive elements made from dielectrics, and also for sensors and energy applications. For example, the modulation of refractive index can be obtained in such components through the change in material composition. Another possibility is to apply concept of gradation in semiconductor active devices. In semiconductors the material function can describe energetic bandgap, refractive index, carrier concentration, carrier mobility, diffusion length, built-in electric field and another property which influence the parameters of optoelectronic devices.

B. High-Efficient Photodetector and Solar Cells [5]

The fundamental limitation of the efficiency of homogenous silicon solar cells is the constant energetic band-gap width in bulk material. Because the high-energetic radiation is absorbed in a shallow layer under surface, it is necessary to form electric field in close vicinity to the surface. Generated carriers can effectively be separated in electric field; therefore the diffusion length of carriers should be longer than junction depth.

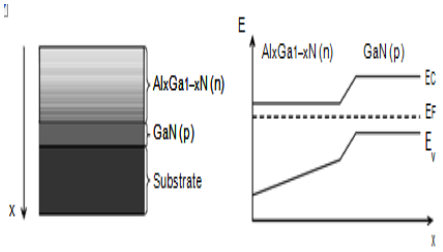


Fig. 3. Schematic structure of *p-n* photodiode with graded layerelectric field, therefore the diffusion length of carriers should be longer than junction depth [3].

Another factor which decreases carrier generation efficiency is the difference of energetic band-gap and absorbed photons energy. By using materials with gradation of energetic band-gap, it is possible to match the absorption edge with band-gap, which improves generation efficiency. The appliance of cascade of junctions with different energetic band-gap width can be one of the solutions. Another way to overcome this limitation is the use of graded material. The idea of such device is shown in Fig. 3.

IV. FGM Thermoelectric Materials

A good thermoelectric material possesses a large Seebeck coefficient, high electrical conductivity and low thermal conductivity. A high electrical conductivity is necessary to minimize the Joule heating, while a low thermal conductivity helps to retain heat at the junctions and maintain a large temperature gradient. These materials, however, have a constant ratio of electrical to thermal conductivity (Widemann-Franz-Lorenz law); so it is not possible to increase one without increasing the other. Metals best suited to thermoelectric applications therefore posses a high Seebeck coefficient. Unfortunately, most of these metals possess Seebeck coefficients in the order of 10 microvolts/K, resulting in generating efficiencies of only fractions of a percent. Therefore, the development of formulated semiconductors with Seebeck coefficients in excess of 100 microvolts/K increased the interest in thermoelectricity. Earlier it was not known that the semiconductors were superior thermoelectric materials due to their higher ratio of electrical conductivity to thermal conductivity, when compared to metals.

A higher figure-of-merit Z of a thermoelectric material shows a higher performance at a specific narrow temperature range [4]. On the other hand, the specific temperature can be shifted to higher temperature by increasing the carrier concentration. Bismuth telluride (Bi_2Te_3), lead telluride (PbTe) and Si-Ge alloy (SiGe) are used for low, medium and high temperature range, respectively. Usually, a monolithic and uniform thermoelectric material is used, though a temperature gradient exists in the thermoelectric material. Therefore, each part has not proper carrier concentration for each temperature. Two times of higher performance than a traditional thermoelectric material can be expected, if the proper carrier concentration gradient is performed to fit with the temperature gradient. Performing stepwise change of carrier concentration is also a performing method for practical application. That is a fundamental concept of energy converting FGM.

It is essential to choose a proper material for each part to fit the temperature gradient. The proper material is a material with proper carrier concentration and a proper compound to match the temperature of each part along the temperature gradient. FGM joining of these materials and fitting electrodes with FGM interface are also core technique, because thermal stress relaxation caused by the difference of thermal expansion coefficient is important at a high temperature.

A. PbTe Based FGM Thermoelectric Material [6]

Performing higher Z over a wide temperature range is essential to obtain higher thermoelectric efficiency η. Figure 4 shows the temperature dependence of Z for 5 kinds of n-type PbTe with different carrier concentration, which was induced from a report of ZT-data by Goff and Lowney as the parameter of carrier concentration. Every Z-curve has a maximum value, and the corresponding temperature shifts with the carrier concentration. If these materials are joined to fit with the temperature gradient, a higher performance than traditional thermoelectric materials can be expected.

When the FGM structure is applied to PbTe, the conversion efficiency can be improved remarkably. Five kinds of PbTe in Fig 4 are joined in sequence at the intercept temperatures of 530, 640, 750 and 815K. This segmented PbTe has five kinds of different carrier concentration and shows five maximum values of Z at different temperatures.

Comparing with the ordinal monolithic PbTe with a carrier concentration of $3x10^{25}/m^3$, which shows the highest efficiency in monolithic PbTe, the FGM exhibits a larger Z value as shown in the hatched area. The broken line in Fig. 4

shows estimation on the assumption of continuous change of the carrier concentration for PbTe. This gradient PbTe has an ideal FGM structure and exhibits a higher *Z* than the segmented PbTe over a wide temperature range. It is estimated that the average Z of the ideal FGM should improve by 50% in comparison with a monolithic PbTe, and maximum thermoelectric efficiency η at T_h=950K should extend to 19%.

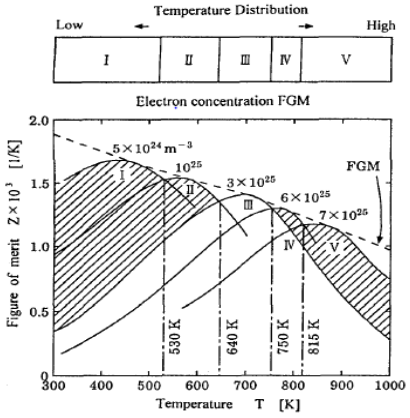


Fig. 4. Temperature dependence of Z for n-type PbTe as parameters of camer concentration[4].

This concept can be applied not only to the carrier concentration gradient but also to joining different materials, such as Bi_2Te_3 -PbTe-SiGe etc., corresponding to the temperature gradient.

V. Conclusions

1) FGM belongs to novel material category and offers new capabilities to use it at large scale. FGM technology has the potential to drastically redefine the methods used for developing lighter, stronger, and high performance structures with unique and nontraditional properties. FGM-composites often lead to a reduction in weight and costs, and are more environmental friendly. For these reasons the popularity of these composites is increasing in engineering world. However, engineering breakthroughs will be required to develop the technology in long-term goals. Technologies expect the FGM kind of advanced materials with revolutionary new capabilities such as in power systems, electronic devices, sensors, science and medical diagnostic instruments.
2) Functionally graded materials are perspective materials for modern optoelectronic devices, such as low threshold current edge lasers (GRINSCH) and tuneable photo detectors. Graded layers can also be used as buffers in heteroepitaxy of nitrides. Development and characterization of FGM is a real challenge and needs thorough analysis. The methodology of measuring layers with graded changing properties is difficult. Another problem is the interrelation between different material properties. It is different to change one parameter like energetic band-gap, without influencing the others like refractive index.
3) The thermoelectric materials with FGM structure show a higher performance than monolithic materials. FGM joining is also a useful technique for setting an electrode in order to relax the thermal stress and suppress the inter diffusion. However, it is difficult to improve the performance by controlling crystal size. On the other hand, exact measurement of thermal conductivity at a high temperature is another problem for accurate designing of an FGM.
4) A temperature sensor results if one of the junctions is maintained at a well-known temperature. The thermocouple can also be used as an actuator. The advantages of TE energy conversion is that moving mechanical parts are avoided which enables high system reliability, quiet operation, and it is usually environmentally friendly. It has been found that the tellurium compounds Bi_2Te_3 and Sb_2Te_3 having ZT > 2 are well established TE materials for futuristic uses.

References

[1] K. M. Gupta, *Materials Science, Metallurgy and Engineering Materials*, Umesh Publications, New Delhi, 2012.
[2] G. Udupal, S. S. Rao, and K. V. Gangadharan, "Future applications of Carbon Nanotube reinforced Functionally Graded Composite Materials," in *Proc. IEEE-International Conference on Advances in Engineering, Science and Management (ICAESM -2012)*, March 30-31, 2012.
[3] A. S. Maiolo, M. N. Amado, and J. S. Gonzalez *et al*., "Development and characterization of poly (vinyl alcohol) based hydrogels for potential use as an articular cartilage replacement," *Mater. Sci. Eng. C*, 2012, vol. 32, no. 6, pp. 1490–1495.
[4] G. Abedi, A. Sotoudeh, M. Soleymani *et al*., "A collagenpoly(vinyl alcohol) nanofiber scaffold for cartilage repair," *J. Biomater. Sci., Polym*. Ed., 2011, vol. 22, no. 18, pp. 2445–2455.
[5] M. Wosko, B. Paszkiewicz, T. Piasecki, A. Szyszka, R. Paszkiewicz, and M. Tlaczala, "Application of functionally graded materials in optoelectronic devices," *optical Application*, vol. 35, no. 3, 2005.
[6] I. Shiota and I. A. Nishida, "Development of FGM thermoelectric materials in Japan-the state of the art," *16th International Conference on Thermoelectrics*, IEEE, pp. 364-370, 1997.

Isha Bharti is from New Delhi, India. She completed her Masters (M.Tech.) in Nanoscience and Technology with specialization in Nanobiotechnology from Delhi Technological University in 2011. Earlier, she did her under graduation (B.Tech.) in Biotechnology from Maharshi Dayanand University, Rohtak, Haryana. Her major field of study is synthesis and fabrication of Nanobiotechnological materials, luminescent materials for novel biomedical applications and diagnostic therapies, etc.

Observational constraints on the kinematics and the fate of the Universe through linearly varying deceleration parameter laws

Özgür Akarsu* and Tekin Dereli†
Department of Physics, Koç University, 34450 Sarıyer, İstanbul, Turkey.

Suresh Kumar‡
Department of Applied Mathematics, Delhi Technological University, Delhi-110 042, India.

Lixin Xu§
Institute of Theoretical Physics, Dalian University of Technology, Dalian, 116024, P. R. China.

(Dated: May 23, 2013)

The deceleration parameter being poorly known by observations but a key parameter in cosmology has been frequently investigated with observational data through its model independent phenomenological parametrizations in terms of cosmic redshift or scale factor. In this paper, we study linearly varying deceleration parameter in terms of cosmic time t (LVDPt) with the companion linearly varying deceleration parameter in terms of cosmic redshift z (LVDPz) and in terms of cosmic scale factor a (LVDPa). We investigate in detail the kinematics and the fate of the Universe by confronting the three LVDP laws with the latest observational data from $H(z)$ compilation (25 data points) and SN Ia Union2.1 compilation (580 data points). The study reveals that the LVDPt law is superior than LVDPz and LVDPa laws in many aspects. In particular, the goodness of fit to the observational data is found to be the best for the LVDPt law. The kinematics and dynamics (assuming general relativity) of the Universe is further studied by considering the LVDPt law in comparison with the standard ΛCDM model. It is found that these two models are observationally indistinguishable but the LVDPt fits the data slightly better than the ΛCDM model. These two models exhibit a very similar behavior for a long passage of time but make slightly different predictions for the earlier times of the Universe with considerably different predictions on the future of the Universe. According to the LVDPt model, the accelerating expansion of the Universe started when the Universe was 7.148 ± 0.923 Gyr old at redshift $0.703^{+0.356}_{-0.148}$ and the age of the present Universe is 13.460 ± 0.899 Gyr, whereas these values are found to be 7.233 ± 0.225 Gyr, 0.682 ± 0.082 and 13.389 ± 0.289 Gyr in the ΛCDM model. Observational constraints on cosmic doomsday are carried out in the context of the LVDPt model, and it is found that the Universe will end in Big Rip when the age of the Universe is 36.602 ± 7.930 Gyr.

Keywords: Accelerating Universe · Linearly varying deceleration parameter · Observational constraints · Cosmic doomsday

I. INTRODUCTION

It was thought that the expansion of the current Universe could be well described by solving Einstein's field equations in the presence of pressure-less matter (dust) within the framework of spatially flat Robertson-Walker spacetime, which gives a decelerating expansion rate with a constant deceleration parameter (DP) equal to $\frac{1}{2}$. The experimental efforts to confirm this, however, led to the discovery that the current Universe is in fact accelerating [1, 2]. It is today very well established that we are living in a Universe that is not only expanding but also have been evolving from decelerating expansion to accelerating expansion and is in the accelerating expansion phase during the last ~ 6 billion years. However, we still have no satisfactory explanation for this fact that occurs at energy scales $\sim 10^{-4}$ eV where we supposedly know physics very well. This accelerated expansion of the Universe, essentially, requires either the presence of an energy source in the context of Einstein's general theory of relativity whose energy density is constant/varies very slowly as the Universe expands and that permeates all over the space uniformly (dubbed as dark energy) [3, 4] or a modification of the Einstein's general theory of relativity for describing gravitation at cosmological scales [5–7]. It seems that the quest for a satisfactory explanation to the accelerating expansion of the current Universe and its history more accurately and precisely from observations will continue to be among the hottest topics in cosmology studies in the near future.

What we expect from a successful cosmological model that provides a possible explanation of the current acceleration of the Universe is not only a successful fit to the observational data but also a motivation from fundamental physics and simplicity. In this respect, so far the best explanation might have been the ΛCDM model that is obtained simply by the inclusion of a positive cosmological constant Λ (which is mathematically

* oakarsu@ku.edu.tr
† tdereli@ku.edu.tr
‡ sukuyd@gmail.com (corresponding author)
§ lxxu@dlut.edu.cn

equivalent to the conventional vacuum energy density that is described with an equation of state (EoS) parameter equal to -1) into the Friedmann equations for spatially flat Universe containing dust [8]. However, it suffers from two serious conceptual/theoretical problems known as the fine tuning and coincidence problems [9–11]. Besides these, the latest cosmic microwave background (CMB) experiment, the *Planck* experiment whose major goal is to test ΛCDM model to high precision and identify areas of tension, concludes that the *Planck* data are remarkably consistent with the predictions of the base ΛCDM model but presence of a mismatch with the temperature spectrum and some other anomalies at low multipoles $l \lesssim 30$ are possibly indicative that the model is either incomplete or alternatively is incorrect [12]. The value of the Hubble constant H_0, which is tightly constrained by CMB data alone in the base ΛCDM model, is found to be strikingly low $H_0 = 67 \pm 1.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in *Planck* experiment. However, the constraints on the kinematics of the Universe from CMB data are highly dependent on the assumed cosmological model and even small changes in the model assumptions, say small deviations from the base ΛCDM model, can change the obtained value of H_0 noticeably. It is, therefore, important to compare the results obtained from CMB data with astrophysical measurements of H_0, since any discrepancies could be a pointer to new physics, particularly to a cosmological model other than the ΛCDM model. For instance, there is a tension at about 2.5σ level between *Planck* and the direct measurements of Hubble constant ($H_0 = 73 \pm 2.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ which is the value obtained using Hubble Space Telescope (HST) observations of Cepheid variables in the host galaxies of eight type Ia supernovae (SN Ia) to calibrate the supernova magnitude-redshift relation [13]) that cannot be resolved by varying the parameters of the ΛCDM model. There is also a mild tension at about 2σ level with the Supernova Legacy Survey (SNLS) compilation while it is compatible with the Union2.1 compilation whose scatter, however, is significantly larger than that of the SNLS compilation. *Planck* experiment concludes that the tension between CMB-based estimates and the astrophysical measurements of H_0 is intriguing and merits further discussion. In fact the supplementary information from astrophysical data sets have an important role (e.g., prevent degeneracies) in the analysis of CMB experiments in general. CMB data alone provide only weak constraints on the dark energy EoS parameter. For instance, again in the *Planck* experiments, while the CMB data alone is compatible with cosmological constant assumption in the base ΛCDM model, the addition of astrophysical data into the analysis usually draws the EoS of the dark energy into the phantom domain ($w < -1$). In such a situation, i.e. if $w \neq -1$, it makes physically sense to think of possible time dependence of w. Accordingly, the *Planck* group considers the simple linear relation $w(a) = w_0 + w_a(1-a)$ (which has often usage in the literature) and conclude that the dynamical dark energy is favored at about 2σ level when they consider the direct measurement of H_0, or the SNLS SNe sample, together with *Planck* data.

A large number of cosmological models have already been proposed as alternatives to the ΛCDM model starting right after the discovery of the current acceleration of the Universe, particularly due to the so called cosmological constant problem. It seems now that the recent results from *Planck* experiment will also not end the inquiries for alternative cosmologies but will lead to quest for cosmological models that does not deviate from the ΛCDM model since the redshift value ~ 3400 and that fit the observational data even better and may relieve the tensions between the different observational data groups.

In the analyses of cosmological data it is usually assumed that general relativity is valid at cosmological scales, the physical ingredient of the current Universe is made of a dark energy source and dust or that a modified gravity is valid at cosmological scales so that an additional dark energy source will not be necessary. Such analyses are often referred to as *dynamical studies*. One of the advantages of such studies is that the cosmological model that will be confronted with the observations can usually be constructed starting from a well motivated theoretical background. However, what is done basically in such studies is to obtain the free parameters of the considered cosmological model using the observational data and the different cosmological models usually give different results for the kinematics of the Universe even in the range covered by the considered observational data and can predict completely different histories for the Universe out of the range covered by the considered observational data. It is, therefore, clear that the expansion history of the Universe obtained in such studies are restricted with the cosmological models that would have been constructed so far. Besides that, we are neither sure that the general relativity needs to be modified nor that the presence of an unknown dark energy source is inevitable. Hence, it makes sense to try to obtain information on the kinematics of the Universe with no assumption for a particular metric theory of gravity or for the matter-energy content of the observed Universe. Such a route has less chance to shed light on the question whether general relativity is valid at cosmological scales or on the nature of the dark energy. However, this route gives us opportunity to investigate the history of the Universe in a less model dependent way/more directly by being able to consider a larger class of expansion histories so that it is more likely to find an expansion law that can fit the observational data better. On the other hand, this route is in fact not totally divorced from the dynamical studies. The considered parametrization for the kinematics of the Universe can often be converted into a dynamical set of equations by assuming a particular metric theory of gravity (for instance general relativity) and then the properties of the

physical ingredient of the Universe can be investigated. A crucial point when this route is taken up is to develop a reasonable strategy for parametrizing the kinematics of the Universe. Various strategies can be followed depending on some practical, mathematical, phenomenological, theoretical and other aspects [14–22]. For instance, the DP being poorly known by observations but a key parameter in cosmology has been constructed observationally through model independent phenomenological parametrizations [23–27]. For the recent studies in this regard see Refs. [28, 29]. It is worth mentioning here that in these studies only cosmic redshift or scale factor-dependent parametrizations of DP have been considered to confront with observations. But in this paper, for the first time in the literature (to the best of our knowledge) we consider a time-dependent parametrization of the DP for its observational construction. This parametrization consists of first two terms of Taylor expansion of DP in cosmic time and hence is naturally motivated, which we refer to as linearly varying DP in cosmic time with the abbreviation LVDPt (See Section II for detailed motivation), which was first introduced by Akarsu and Dereli [30, 31]. We also consider the companion linearly varying DP in terms of cosmic redshift z (LVDPz) [23, 24, 27, 28] and in terms of cosmic scale factor a (LVDPa)[24–26, 28] for the investigation in contrast with LVDPt.

In the light of the above discussion, here we study in detail the kinematics and the fate of the Universe by confronting the three LVDP laws with the latest observational data from $H(z)$ compilation (25 data points) and SN Ia Union2.1 compilation (580 data points). The paper is organized as follows. In Section II, we first discuss our strategy which leads to the three different LVDP laws, viz., LVDPz, LVDPa and LVDPt. We then discuss some properties of these three LVDP laws from various aspects and compare to each other. In Section III, we obtain the observational constraints on the kinematics of the Universe from the latest $H(z)$+SN Ia data (25+580 data points) using the three LVDP laws separately. In Section IV, we further discuss the constraints on the dynamics of the Universe using the LVDPt law in comparison with the ΛCDM model as well as with some other models such as DGP, Chaplygin gas, Galileon models. In Section V, we discuss the predicted future of the Universe using the laws we considered in the paper, particularly using LVDPt law. Final section of the paper is filled with the concluding remarks.

II. THE MOTIVATION FOR LVDP LAWS AND THEIR THEORETICAL COMPARISONS

Three important parameters in cosmology are the Hubble parameter H, dimensionless DP q and the dimensionless jerk parameter j that are defined as follows:

$$H=\frac{\dot{a}}{a},\quad q=-\frac{\ddot{a}a}{\dot{a}^2}\quad\text{and}\quad j=\frac{\dddot{a}a^2}{\dot{a}^3}, \tag{1}$$

where a is the scale factor and a dot denotes $\frac{\mathrm{d}}{\mathrm{d}t}$. DP among these is the key quantity in describing the evolution of the Universe. Assuming general relativity is the true theory of gravitation, it also tells us one of the most important parameter for understanding the physical ingredient of the Universe, the EoS parameter $w=p/\rho$ (where ρ and p are the energy density and pressures respectively) of the effective fluid that controls the dynamics of the Universe, through the relation

$$q=\frac{1}{2}+\frac{3}{2}w, \tag{2}$$

in the framework of spatially flat RW spacetime. Accordingly, before the first indications of the current acceleration of the Universe, in 1998, the DP of the Universe was expected to be $q\cong\frac{1}{2}$ since it was believed that the physical ingredient of the current Universe could be described very well with a dust with a negligible EoS parameter $w\cong 0$.

We would like to show here what we get if constant DP (CDP) $q=\frac{1}{2}$ prediction is relaxed to the general CDP $q=q_0$ and constrain the corresponding cosmologies $H(z)=H_0(1+z)^{\frac{3}{2}}$ and $H(z)=H_0(1+z)^{1+q_0}$, respectively, with the latest data from $H(z)$ (25 data points) and SN Ia observations (580 data points) (the details on the constraints methodology are given in Section III). We give the results in Table I. Observational constraints on H_0 and q_0 in the cosmology $H(z)=H_0(1+z)^{1+q_0}$ (power-law cosmology) were carried out recently by Kumar [32], separately and also jointly, using 14 observational $H(z)$ data points from [13, 33] and from Union2 set of 557 SN Ia data points [34]. In that study, from the joint analysis, the Hubble constant and DP were found as $H_0=68.93^{+0.53}_{-0.52}$ and $q_0=-0.34\pm0.05$. The different values we obtain for CDP (given in Table I) are due to the updated data (25 $H(z)$ and 580 SN Ia data points) that we have considered in this study (See Section III).

TABLE I: Mean values with 1σ, 2σ and 3σ errors of the parameters of CDP and CDP($q=\frac{1}{2}$) models constrained with $H(z)$+SN Ia data. H_0 is given in the units $\mathrm{km\,s^{-1}\,Mpc^{-1}}$

	CDP	CDP($q=\frac{1}{2}$)
q_0	$-0.192^{+0.063+0.118+0.174}_{-0.064-0.126-0.170}$	0.5
H_0	$67.716^{+0.832+4.181+5.148}_{-1.151-3.420-5.687}$	$53.828^{+1.318+16.752+17.452}_{-2.405-3.571-4.971}$
χ^2_{min}	567.413	719.438

We note first that the value of the χ^2_{min}, i.e., the goodness of fit, improved considerably when $q=\frac{1}{2}$ is relaxed

to be $q = q_0$. The value $q = \frac{1}{2}$ is not within the error region of the one we obtain when it is relaxed even within 3σ error region. These results can be interpreted mainly in two different ways: (i) If we assume that general relativity is valid, then the Universe in the range covered by the data cannot be described with $w \cong 0$ ($q \cong \frac{1}{2}$) on average but with $w \cong -0.14$ ($q \cong -0.192$), which indicates the presence of an unknown fluid with a negative EoS parameter in addition to the dust. (ii) If we assume that $w \cong 0$ is a valid assumption for describing the physical ingredient of the Universe in the range covered by the data then the general relativity could not be valid at cosmological scales. In the standard ΛCDM model, on the other hand, what is constant is not the DP but the jerk parameter: jerk parameter is simply a constant equal to unity $j_{\Lambda\mathrm{CDM}} = 1$, while the DP evolves to -1 starting from $\frac{1}{2}$. However, compared to the Hubble constant and DP, the observational constraints on the value of the jerk parameter are rather weak since various observational studies give values approximately in the range $-5 \lesssim j_0 \lesssim 10$ [15, 35–41]. Therefore, the purpose of the current cosmological studies is primarily an accurate and precise description of the evolution of DP and then to examine the behavior of the jerk parameter in comparison with $j_{\Lambda\mathrm{CDM}} = 1$ of the ΛCDM model.

The evolution of DP can, in principle, be described by functions of some other cosmic parameters such as cosmic time t, cosmic scale factor a, cosmic redshift z, which can also be re-written in terms of each other. However, we don't know the accurate function that describes its evolution. Let us now make a very reasonable assumption that the unknown function describing the DP can be approximated by its Taylor expansion in terms of x:

$$q(x) = q_0 + q_1\left(1 - \frac{x}{x_0}\right) + q_2\left(1 - \frac{x}{x_0}\right)^2 + ..., \quad (3)$$

where x is a reasonable cosmological parameter to be decided such as cosmic red-shift z, cosmic scale factor a, cosmic time t and subscript $_0$ stands for the current value of the considered parameter. It is obvious that the first approximation to the unknown function $q(x)$ gives a CDP that corresponds to either power-law or exponential expansion, and is independent of the choice of the cosmological parameter for Taylor expansion. However, for the higher order approximations, one should give a decision on a reasonable cosmological parameter (e.g., t, a, z) for the Taylor expansion and on the number of terms to be expanded. The second approximation in terms of x to the unknown function of the DP will be a *linearly varying DP* (LVDP), the third approximation will be a *quadratically varying DP* (QVDP) and so on. We emphasize that each of these decisions will correspond to a different cosmological model/expansion history of the Universe! In other words, even if we obtain similar kinematics in the range covered by the considered observational data, the prediction of each model could be quite different out of this range. For instance, linear approximation for DP in cosmic redshift z is not a linear approximation in cosmic scale factor a and these two linear approximations are obviously quite different since z diverges as a goes to zero. Hence, one may easily conclude also that different choices (such as z, a, t and etc.) for the parameter x for the Taylor expansion and the number of terms considered would fit the observational data differently. In other words, the goodness of the fit will be related with (i) the chosen parameter (t, a, z) for the Taylor expansion and (ii) the number of terms considered. However, in physics, we should consider not only the goodness of fit but also the simplicity of the model (namely, number of terms in the Taylor expansion) in accordance with the Occam's razor principle. If too many terms are considered then the allowed region in the parameter space could be so large that it would not be possible to get firm conclusions. If not enough number of terms are considered, on the other hand, then the observational data would be evaluated on average and misleading conclusions could be reached. Hence we should build a balance between the goodness of the best fit to the observational data and the simplicity of the model to be fitted to the observational data. It is obvious that, for a chosen parameter (t, a, z) for the Taylor expansion, one would obtain better fit to the observational data when the number of terms of Taylor expansion is increased. However, this would yield a more complex model/a model with more free parameters. On the other hand, goodness of the fit to the observational data can also be improved by changing the cosmic parameter we consider for the Taylor expansion instead of increasing the number of terms for Taylor expansion, which leads to more complex models.

In the light of the above discussion, we consider the truncation at the second order,

$$q(x) = q_0 + q_1\left(1 - \frac{x}{x_0}\right), \quad (4)$$

and utilize it, in what follows, to obtain the DP laws linearly varying with z, a and t.

The first LVDP law we shall consider is the LVDP in terms of cosmic redshift z (LVDPz), which has been the first and most commonly used law in the literature for obtaining the kinematical parameters of the Universe using $H(z)$ and/or SNe Ia data [23, 24, 27, 28]. To get LVDPz, we simply substitute $\frac{x}{x_0} = \frac{a_0}{a} = 1 + z$ in (4) and absorb the minus sign with q_1:

$$q = q_0 + q_1 z, \quad (\mathrm{LVDP}z) \quad (5)$$

where $z = -1 + \frac{a_0}{a}$ is the cosmic redshift. The most obvious shortcoming of the LVDPz law is that it grows monotonically with no limits as we go back to earlier times of the Universe. This would make the results obtained by this law unreliable particularly if the observations under consideration involve high redshift data. For instance, let us assume that the current value of the DP is $q_{z=0} \sim -0.5$ and the acceleration started at

$z \sim 0.5$, which implies $q_1 \sim 0.5$ in (5). We note that, in this case, $q = 1$ is reached when $z = 3$, which obviously cannot be accepted since Big Bang nucleosynthesis constrains DP to be $q \sim 1$ even when $z \sim 10^9$ [42]. This law, on the other hand, is well behaved in the future such that the Universe either expands forever (provided that $q_0 \geq q_1 - 1$) or ends with a Big Rip in finite future (provided that $q_0 < q_1 - 1$). However, because the observational constraints obtained using this law with high redshift data are not reliable, the predicted future of the Universe using this law cannot be reliable. Another shortcoming of this law is that we don't have analytical expression of the scale factor, Hubble parameter etc. in terms of cosmic time t for this law.

Unlimited growth of the DP for large z values in the LVDPz law led some authors [24–26, 28] to propose linearly varying DP in scale factor a (LVDPa), which can be obtained, in our study, by substituting $\frac{x}{x_0} = \frac{a}{a_0}$ in (4):

$$q = q_0 + q_1 \left(1 - \frac{a}{a_0}\right). \quad \text{(LVDP}a\text{)} \tag{6}$$

It can also written in terms of redshift as follows

$$q = q_0 + q_1 \frac{z}{1+z} \cong q_0 + q_1 z - q_1 \, (z^2 - z^3 + ...). \tag{7}$$

At this stage, it should be noted that in all the LVDP laws the transition of the Universe from deceleration to acceleration demands the constant q_1 to be positive. One may observe that LVDPz and LVDPa laws exhibit similar behaviors at low redshift values. On the other hand, LVDPa law gets finite values in the whole past of the Universe, which makes this law more reliable than the LVDPz law particularly for observational studies involving high redshift data. Nevertheless, this time we cannot rely on the predictions about the future of the Universe since the LVDPa grows very rapidly, such that the DP $q \to -\infty$ (hence the EoS parameter of the effective fluid $w \to -\infty$ when the general relativity is considered) as $z \to -1$, which we will call **Super Big Rip** and cannot be taken as a realistic prediction. Besides this extreme behavior, this law suffers from the same problem with the LVDPz law such that we don't have analytical expressions of the cosmological parameters in terms of cosmic time t.

LVDPz and LVDPa laws are the most considered two laws in the literature [23–28] to obtain the observational constrains on the kinematics of the Universe and correspond to Taylor expansions of the unknown function of the DP up to the second term in terms of cosmic redshift z and a, respectively. However, in fact, the most essential concept when we are dealing with a dynamical system is time and hence is the cosmic time t in cosmology. Similar to the LVDPz and LVDPa laws, Taylor expansion of the unknown function of the DP in terms of t up to the second term will give a linearly varying DP in time t (LVDPt), which was recently suggested by Akarsu and Dereli [30, 31] phenomenologically. LVDPt law can be obtained simply by substituting $\frac{x}{x_0} = \frac{t}{t_0}$, where t_0 is the present time of the Universe, into (4):

$$q = q_0 + q_1 \left(1 - \frac{t}{t_0}\right). \quad \text{(LVDP}t\text{)} \tag{8}$$

This law can also be re-written and Taylor expanded in terms of redshift as follows:

$$\begin{aligned} q &= q_0 + q_1 \left(1 - \frac{2 + 2q_0 + 2q_1}{q_1 + (2 + 2q_0 + q_1)(1+z)^{1+q_0+q_1}}\right) \\ &\cong q_0 + \frac{1}{2}(2 + 2q_0 + q_1)q_1 z - \frac{1}{4}(2 + q_0)(2 + 2q_0 + q_1)q_1 z^2 + ... \end{aligned} \tag{9}$$

LVDPt law has many advantage compared to the LVDPz and LVDPa laws. We note that at low redshift values all the three LVDP laws behave similarly while only the LVDPt law never diverges through the history of the Universe, namely, it is finite both at the beginning $a \to 0$ ($z \to \infty$) and at the end $a \to \infty$ ($z \to -1$) of the Universe. While we don't have analytical expressions of the cosmological parameters in terms of cosmic time t for the LVDPz and LVDPa laws, we have this law already in terms of cosmic time t and can obtain all the cosmological parameters (a, H, q, j and etc.) in terms of t, a and z whenever it is needed. Moreover, as it was shown in [31], it exhibits a very similar behavior with the ΛCDM cosmology for a long passage of time. This is important since although we are not sure today that ΛCDM is the right model to explain the observed Universe, we know that any cosmological model that can be a good candidate cannot deviate from ΛCDM kinematics a lot in the observational range of the Universe. We note that Universe inevitably evolves into a super-exponential expansion phase with a finite DP value less than -1 that imposes a Big Rip end to the Universe [43], which also imposes, in general relativity, the presence of dark energy source that can cross below the phantom divide line ($w < -1$) that is dubbed as quintom dark energy [44, 45]. This is a behavior that is favored persistently by many observational studies [12, 46–50]. We would also like to comment that LVDPt behavior has recently been obtained by Akarsu and Dereli [51] in a cosmological model in which general relativity and

dark energy source are not considered but the presence of higher dimensions and dilaton gravity as the theory of gravitation.

We compare these three different linear parametrizations of the deceleration parameter, LVDPt, LVDPa and LVDPz in Table II. The constraints on the kinematics of the Universe from observational data using the LVDPt law has not been studied before, whereas the LVDPz and LVDPa laws have been considered in many studies [23–28]. The reason may be as follows: The studies on the kinematics of the Universe are basically based on the redshift concept since the $H(z)$ data consist of the Hubble parameter versus the redshift and the SN Ia data consist of the distance versus redshift. In addition to this, the conversion between the redshift z and scale factor a is rather trivial, i.e. $z = -1 + a_0/a$. Hence what has been considered often in the first place was a relation between the DP and the redshift or scale factor. On the other hand, if we start with a law for DP in terms of t, then we should first obtain the scale factor explicitly to derive the Hubble parameter in terms of redshift so that we will use it in our observational analyses. Indeed, one may observe from Table II that the expression of the LVDPt law is quite non-trivial either in terms z or a, while it is quite trivial to write LVDPa law in terms of z and LVDPz law in terms of a.

III. OBSERVATIONAL CONSTRAINTS ON LVDP LAWS FROM $H(z)$+SN IA DATA

The parameters of the LVDP laws are constrained considering the 25 observational $H(z)$ data points spanning the redshift range $0.07 < z < 1.750$ [52] and the SNe Ia Union2.1 sample [53] that contains 580 SNe Ia data points spanning the redshift range $0.015 < z < 1.414$ and using the Markov Chain Monte Carlo (MCMC) method, whose code is based on the publicly available package **cosmoMC** [54]. The details of methodologies and observational data points used for obtaining the constraints on model parameters are given in Appendices A and B. We combine $H(z)$ and SNe Ia data in order to obtain tighter constraints on the model parameters and to avoid degeneracy in the observational data. Since independent cosmological probes provide the $H(z)$ and SNe Ia data, it is reasonable to define the total likelihood as the product of separate likelihoods of the two probes, viz.,

$$\chi^2_{\rm total} = \chi^2_{OHD} + \chi^2_{\rm SN}.$$

The Hubble parameters in terms of cosmic redshift z, that are used for fitting the data, obtained from the LVDPz, LVDPa and LVDPt laws, respectively, read as

$$H(z) = H_0(1+z)^{1+q_0-q_1}e^{q_1 z}, \quad (\text{LVDP}z) \tag{10}$$

$$H(z) = H_0(1+z)^{1+q_0+q_1}e^{-q_1\frac{z}{1+z}}, \quad (\text{LVDP}a) \tag{11}$$

$$H(z) = \frac{H_0}{4(1+q_0+q_1)^2}\left[q_1(1+z)^{-(1+q_0+q_1)/2} + (2+2q_0+q_1)(1+z)^{(1+q_0+q_1)/2}\right]^2. \quad (\text{LVDP}t) \tag{12}$$

The 1D marginalized distribution on individual parameters and 2D contours with 68.3 %, 95.4 % and 99.73 % confidence limits are shown in Fig.1a for the LVDPt law, in Fig.1b for the LVDPz law and in Fig.1c for the LVDPa law. The mean values of the model parameters q_0 (the current value of the DP), q_1 and the Hubble constant H_0 (the current value of the Hubble parameter) of the LVDPt, LVDPa and LVDPz constrained with $H(z)$+SN Ia data are given in Table III. The 1σ errors, χ^2_{min}, χ^2_{min}/dof and goodness of fit (GoF) are also given in Table III.

We note that the LVDPt law fits the observational data better than the LVDPa and LVDPz laws. LVDPa, on the other hand, fits the observational data better than the LVDPz law.

In Table IV, we give the values of the cosmological parameters that we obtained using the three LVDP laws separately; age of the present Universe t_0; Hubble constant H_0, current value of the DP q_0 and jerk parameter j_0, age of the Universe at the onset of accelerating expansion $t_{\rm tr}$, time passed since the accelerating expansion started $t_0 - t_{\rm tr}$ and redshift of the onset of the accelerating expansion $z_{\rm tr}$.

We note that all the values obtained using LVDPt and LVDPa laws are consistent within 1σ error region. The Hubble constant values are consistent within 1σ error region. in the three LVDP laws. However, the age of the present Universe obtained in LVDPz has a too low value (11.934 ± 0.535 Gyr) and is neither consistent with the one obtained in LVDPt law nor with the one obtained in LVDPa law within 1σ error region. The current values of the DP obtained in all three LVDP

TABLE II: Comparison of LVDPt, LVDPa and LVDPz laws.

	LVDPt	LVDPa	LVDPz
$q(t)$	$q_0 + q_1(1 - \frac{t}{t_0})$	Not available	Not available
$q(a)$	$q_0 + q_1\left(1 - \frac{2+2q_0+2q_1}{q_1+(2+2q_0+q_1)(a/a_0)^{-1-q_0-q_1}}\right)$	$q_0 + q_1(1 - \frac{a}{a_0})$	$q_0 + q_1\left(-1 + \frac{a_0}{a}\right)$
$q(z)$	$q_0 + q_1\left(1 - \frac{2+2q_0+2q_1}{q_1+(2+2q_0+q_1)(1+z)^{1+q_0+q_1}}\right)$	$q_0 + q_1\frac{z}{1+z}$	$q_0 + q_1 z$
$q(z \to \infty)$	$q_0 + q_1$	$q_0 + q_1$	$+\infty$
$q(z = 0)$	q_0	q_0	q_0
$q(z = -1)$	$-q_0 - q_1 - 2$	$-\infty$	$q_0 - q_1$
Future	Big Rip	Super Big Rip	Big Rip if $q_0 < q_1 - 1$
			Expands forever if $q_0 \geq q_1 - 1$

TABLE III: Mean values with 1σ errors of the parameters of LVDPt, LVDPz and LVDPa models constrained with $H(z)$+SN Ia data. χ^2_{min}/dof and GoF values are also given.

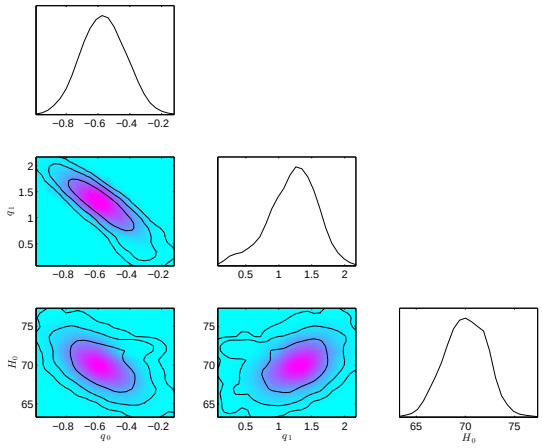
Parameters	LVDPt	LVDPa	LVDPz
q_0	$-0.565^{+0.141}_{-0.135}$	$-0.613^{+0.137}_{-0.141}$	$-0.478^{+0.106}_{-0.107}$
q_1	$1.206^{+0.352}_{-0.344}$	$1.540^{+0.453}_{-0.442}$	$0.652^{+0.194}_{-0.197}$
H_0 (km s^{-1} Mpc^{-1})	$70.016^{+2.089}_{-2.096}$	$70.260^{+1.887}_{-2.102}$	$69.470^{+1.991}_{-1.964}$
χ^2_{min}	556.489	556.508	557.314
χ^2_{min}/dof	0.91981	0.91984	0.92118
GoF	92.137	92.129	91.763

TABLE IV: Mean values with 1σ errors of some important cosmological parameters pertaining to LVDPt, LVDPz and LVDPa laws.

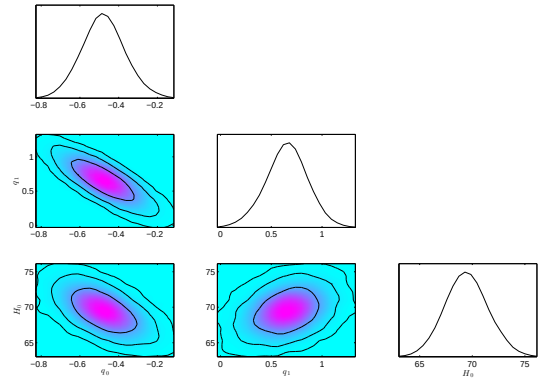
Parameters	LVDPt	LVDPa	LVDPz
t_0 (Gyr)	13.460 ± 0.899	13.138 ± 0.754	11.934 ± 0.535
H_0 (km s^{-1} Mpc^{-1})	$70.016^{+2.089}_{-2.096}$	$70.260^{+1.887}_{-2.102}$	$69.470^{+1.991}_{-1.964}$
q_0	$-0.565^{+0.141}_{-0.135}$	$-0.613^{+0.137}_{-0.141}$	$-0.478^{+0.106}_{-0.107}$
j_0	1.326 ± 0.599	1.680 ± 0.645	0.631 ± 0.290
$t_{\rm tr}$ (Gyr)	7.148 ± 0.923	7.056 ± 0.754	5.430 ± 0.529
$t_0 - t_{\rm tr}$ (Gyr)	6.312 ± 1.288	6.082 ± 1.066	6.504 ± 0.752
$z_{\rm tr}$	$0.703^{+0.356}_{-0.148}$	$0.662^{+0.198}_{-0.107}$	$0.733^{+0.148}_{-0.095}$

laws are consistent within 1σ error region. The current values of the jerk parameter obtained in LVDPt and LVDPa laws are consistent within 1σ error region and also cover $j_{\Lambda\rm CDM} = 1$. However the current value of the jerk parameter obtained using LVDPz law is not only inconsistent with the ones obtained in LVDPt and LVDPa laws but also does not cover $j_{\Lambda\rm CDM} = 1$ within 1σ error region. Considering its poor fit to the data and too low age for the current Universe, we conclude that the LVDPz law is not reliable for obtaining the kinematics of the Universe. According to LVDPt and LVDPa laws, the accelerating expansion of the Universe starts when the Universe was ~ 7 Gyr old (~ 6 Gyr ago from now) and at redshift ~ 0.7. However, considering the goodness of the fits, the results obtained using the LVDPt law can be taken as more reliable than the results obtained using the LVDPa law. This is important also because the predictions for the future and the past of the Universe will be quite different in LVDPt and LVDPa laws though these two laws give similar results within 1σ error region (See Table III and Table IV). Jerk parameter is a key parameter to investigate the deviations from the ΛCDM model in which jerk parameter is simply a constant equal to 1. We note that the present value of the jerk parameter does not deviate from $j = 1$ in all the three LVDP laws but it is covered only by the LVDPt law within 1σ error region.

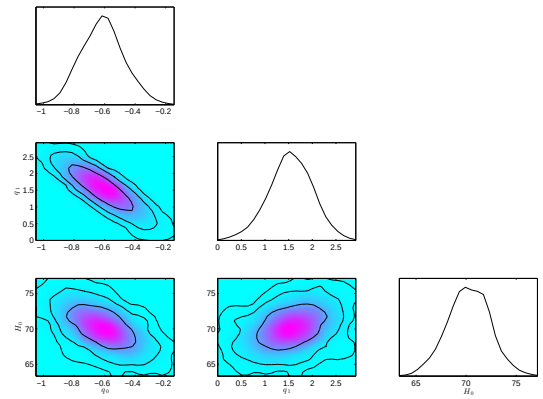
We conclude that it is tempting to compare LVDPt with ΛCDM cosmology since the fit to the observational data provided with the highest success in the LVDPt law and the present value of the jerk parameter is consistent with $j_{\Lambda\rm CDM} = 1$ in this law, which indicates that LVDPt



(a) LVDPt



(b) LVDPz



(c) LVDPa

FIG. 1: The 1D marginalized distribution on individual parameters of (a) LVDPt, (b) LVDPz and (c) LVDPa and 2D contours with 68.3 %, 95.4 % and 99.7 % confidence levels are obtained by using $H(z)$+SN Ia data points. The shaded regions show the mean likelihood of the samples.

and ΛCDM models have very similar behavior at least in the vicinity of the current Universe.

IV. COMPARISON OF THE LVDPt AND ΛCDM DYNAMICS

In this section, considering the observational constraints from the latest $H(z)$ and SN Ia data (See Section III), we compare the dynamics of the Universe model obeying the kinematics of the LVDPt with the ΛCDM cosmology. We compare the effective fluid that would give rise to LVDPt kinematics considering general relativity with that of the ΛCDM model where the effective fluid is a mixture of the conventional vacuum energy and dust.

The scale factor in terms of cosmic time t for the LVDPt law (8) is as follows:

$$a(t) = a_1 e^{\frac{2}{1+q_0+q_1}\tanh^{-1}\left[\frac{q_1 t}{(1+q_0+q_1)t_0}-1\right]}, \qquad (13)$$

where a_1 is a constant of integration. The Hubble parameter

$$H = \frac{2t_0}{t[2(1+q_0+q_1)t_0 - q_1 t]}, \qquad (14)$$

and the dimensional jerk parameter

$$j = \frac{3q_1^2 t^2}{2t_0^2} - \frac{3q_1(q_0+q_1)t}{t_0} + (q_0+q_1)(2q_0+2q_1+1). \qquad (15)$$

In general relativity within the framework of spatially flat Friedmann-Robertson-Walker space-time the effective energy density $\rho = (3/8\pi G)H^2$ and pressure $p = -(c^2/8\pi G)(2\dot{H}+3H^2)$. Using these we obtain the energy density and pressure of the effective fluid that would give rise to the LVDPt kinematics in general relativity as

$$\rho = \frac{3t_0^2}{2\pi G t^2[2(1+q_0+q_1)t_0 - q_1 t]^2}, \qquad (16)$$

$$p = \frac{c^2 t_0[-2q_1 t + (2q_0+2q_1-1)t_0]}{2\pi G t^2[2(1+q_0+q_1)t_0 - q_1 t]^2}, \qquad (17)$$

respectively. Consequently, we obtain the effective EoS parameter as follows:

$$w = -1 + \frac{2(1+q_0+q_1)}{3} - \frac{2q_1 t}{3t_0}. \qquad (18)$$

We see that the parameters H, ρ and p diverge at $t = 0$ and $t = 2(1+q_0+q_1)t_0/q_1$ while q, j and w are finite at these two moments. This in turn implies that the Universe governed by LVDPt law begins with a Big Bang at $t = 0$ and ends in a Big Rip at $t_{\rm BR} = 2(1+q_0+q_1)t_0/q_1$. Thus, the LVDP$t$ Universe has finite

TABLE V: Mean values with 1σ errors of the parameters of LVDPt and ΛCDM models constrained with $H(z)$+SN Ia data. χ^2_{min}/dof and GoF values are also displayed.

Parameters	LVDPt	ΛCDM
q_0	$-0.565^{+0.141}_{-0.135}$	-0.556 ± 0.046
q_1	$1.206^{+0.352}_{-0.344}$	–
H_0 (km s^{-1} Mpc^{-1})	$70.016^{+2.089}_{-2.096}$	$70.697^{+1.667}_{-2.020}$
χ^2_{min}	556.489	556.499
χ^2_{min}/dof	0.91981	0.91983
GoF	92.137	92.133

life time. The de Sitter time that corresponds to $q = -1$ reads as $t_{\rm ds} = (1 + q_0 + q_1)t_0/q_1$. It is interesting to observe the coincidence that the de Sitter time in the LVDPt model is exactly half the Big Rip time, that is, $t_{\rm ds} = \frac{1}{2}t_{\rm BR}$.

Further, we would like to note here that one is able to re-write all these cosmological parameters that are given in terms of cosmic time t in terms cosmic redshift z using the relation

$$t = \frac{2(1 + q_0 + q_1)t_0}{q_1 + (2 + 2q_0 + q_1)(1 + z)^{1+q_0+q_1}}, \qquad (19)$$

between the t and z that can be obtained using $a = 1/(1 + z)$ with the choice $a(t_0) = 1$ for the present size of the Universe.

In Table V, we give the mean values with 1σ errors of the parameters of LVDPt and ΛCDM models constrained with $H(z)$+SN Ia data. The values of χ^2_{min}, χ^2_{min}/dof and GoF are also displayed in Table V. We note that LVDPt model fits the observational data slightly better than the ΛCDM model does.

We give cosmological parameters including the present values of the effective energy densities ρ_0, pressures p_0 and EoS parameters w_0 for both models in Table VI. We observe that the present values of the cosmological parameters as well as the parameters related with the onset of accelerating expansion obtained in LVDPt and ΛCDM are very similar/indistinguishable. Strictly speaking all the values obtained using LVDPt law and ΛCDM are consistent even at 1σ level.

The values we obtained for the Hubble constant from the latest $H(z)$+SNI a Union2.1 compilation are completely consistent even within 1σ error region; $H_0 = 70.016^{+2.089}_{-2.096}\,{\rm km\,s^{-1}\,Mpc^{-1}}$ in the LVDPt model and $H_0 = 70.697^{+1.667}_{-2.020}\,{\rm km\,s^{-1}\,Mpc^{-1}}$ in the ΛCDM model. These values are consistent with the low value $H_0 = 70.0 \pm 2.2\,{\rm km\,s^{-1}\,Mpc^{-1}}$ that has been found in the recent WMAP-9 analysis [55] assuming the base six-parameter ΛCDM model. On the other hand, only the one obtained in LVDPt law is consistent with the low value $H_0 = 67.3 \pm 1.2\,{\rm km\,s^{-1}\,Mpc^{-1}}$ [12] from the Planck+WP+highL analysis within 1σ error region in the *Planck* experiment.

TABLE VI: Mean values with 1σ errors of some important cosmological parameters related to LVDPt and ΛCDM models.

Parameters	LVDPt	ΛCDM
t_0 (Gyr)	13.460 ± 0.899	13.389 ± 0.289
H_0 (km s^{-1} Mpc^{-1})	$70.016^{+2.089}_{-2.096}$	$70.697^{+1.667}_{-2.020}$
q_0	$-0.565^{+0.141}_{-0.135}$	-0.556 ± 0.046
j_0	1.326 ± 0.599	1
$t_{\rm tr}$ (Gyr)	7.148 ± 0.923	7.233 ± 0.225
$t_0 - t_{\rm tr}$ (Gyr)	6.312 ± 1.288	6.156 ± 0.366
$z_{\rm tr}$	$0.703^{+0.356}_{-0.148}$	0.682 ± 0.082
ρ_0 (10^{-27} kg m^{-3})	9.209 ± 0.545	9.389 ± 0.481
p_0 (10^{-10} Pa)	-5.890 ± 0.976	-5.952 ± 0.520
w_0	-0.710 ± 0.090	-0.704 ± 0.030

We expect these two models to differ slightly at earlier times of the Universe and differ significantly in the far future. Therefore, we compare the asymptotic values of various parameters of the LVDPt and ΛCDM models in Table VII. We note that LVDPt and ΛCDM models differ only slightly (such that the results are still consistent within 1σ error region) at the limit $z \to \infty$. In the future, on the other hand, these two models behave totally differently. While the ΛCDM model approaches continuously to the de Sitter expansion without an end, LVDPt model evolves to super exponential expansion rate with a value $q = -2.640 \pm 0.238$ as $z \to -1$ which indicates that the size of the Universe will diverge in finite time. We would like to remind that the ΛCDM model is in fact a model that can describe the evolution of the observed Universe starting from the matter dominance $w \sim 0$ (which gives $q \sim \frac{1}{2}$ in general relativity) at $z \sim 3400$. The LVDPt model can be interpreted, in a similar manner, as a model that can describe the evolution of the Universe starting from the matter dominance in the context of dark energy concept, since the value $q = \frac{1}{2}$ of the ΛCDM as $z \to \infty$ is covered by the LVDPt law. However, LVDPt law may be describing the evolution of the Universe starting from even earlier times of the Universe than the ΛCDM model starts to describe. In the actual Universe, the physical ingredient of the Universe at redshift values $z \gtrsim 3400$ is expected

TABLE VII: Mean values and asymptotic limits with 1σ errors of various parameters pertaining to the LVDPt and ΛCDM models.

Model $\rightarrow$	LVDPt			ΛCDM		
Parameter	$z \to \infty$	$z = 0$	$z \to -1$	$z \to \infty$	$z = 0$	$z \to -1$
H (km s^{-1} Mpc^{-1})	∞	$70.016^{+2.089}_{-2.096}$	∞	∞	$70.697^{+1.667}_{-2.020}$	59.336 ± 2.592
q	0.640 ± 0.238	$-0.565^{+0.141}_{-0.135}$	-2.640 ± 0.238	0.5	-0.556 ± 0.046	-1
j	1.462 ± 0.848	1.326 ± 0.599	11.308 ± 2.277	1	1	1
ρ (10^{-27} kg m^{-3})	∞	9.209 ± 0.545	∞	∞	9.389 ± 0.481	6.614 ± 0.577
p (10^{-10} Pa)	∞	-5.890 ± 0.976	$-\infty$	-5.952 ± 0.520	-5.952 ± 0.520	-5.952 ± 0.520
w	0.093 ± 0.158	-0.710 ± 0.090	-2.093 ± 0.158	0	-0.704 ± 0.030	-1

to be described with an EoS parameter $w \sim \frac{1}{3}$, which gives $q \sim 1$ in general relativity, since radiation would start to be dominant during that epoch as we go back to earlier times of the Universe. We note that these values are almost covered at 1σ level by the predicted values for $z \to \infty$ in the LVDPt model. This may be suggesting that the LVDPt law is a good approximation for describing the evolution of the Universe even starting from the Big Bang Nucleosynthesis times.

We have discussed above the values of the cosmological parameters in LVDPt and ΛCDM models for the present Universe and for two extremes $z \to \infty$ and $z \to -1$. We would like to conclude this section by comparing the continuous evolution of these models instead of some particular times of the Universe. A very useful way of comparing and distinguishing different cosmological models that have similar kinematics is to plot the evolution trajectories of the $\{q, j\}$ and $\{j, s\}$ pairs. Here q and j have the usual meaning and s is a parameter we defined as

$$s = \frac{j - 1}{3(q - 1)} \tag{20}$$

inspired by Sahni et al. [56]. Using (13) in (20) we get

$$s = \frac{3q_1^2t^2 - 6t_0q_1(q_0 + q_1)t + 2t_0^2(1 + q_0 + q_1)(-1 + 2q_0 + 2q_1)}{6t_0[-q_1t + (-1 + q_0 + q_1)t_0]} \tag{21}$$

for the LVDPt law. Note that the parameter s we defined is slightly different from the one given as $s = \frac{j-1}{3(q-\frac{1}{2})}$ by Sahni et al. [56]. In our definition we use 1 in the place of $\frac{1}{2}$ in the original definition to avoid the divergence of the parameter s when model passes through $q = \frac{1}{2}$. The parameter s was originally introduced to characterize the properties of dark energy and hence the evolution of the Universe was considered starting from pressure-less matter dominated era in general relativity which gives $q = \frac{1}{2}$ and $j = 1$. However, in accordance with the LVDPt, here we are also interested in the possibility of describing the Universe starting from times even earlier than the pressure-less matter dominated era and may be starting from primordial nucleosynthesis times where the expansion of the Universe can be best described by $q \sim 1$.

We plot evolution trajectories of the LVDPt and ΛCDM models in the $j - q$ plane in Fig. 2a and in the $j - s$ plane in Fig. 2b, using the mean values of the model parameters given in Table V from observations. For comparison, we include also some alternatives to the ΛCDM model such as the Galileon, Chaplygin Gas and DGP models (for these models see [57] and references therein) in the figures. The arrows on the curves show the direction of evolution and the dots on the curves represent the present values of the corresponding $\{q, j\}$ and $\{j, s\}$ pairs while the black dots show the matter dominated phases of the models.

We observe that all the models have different evolution trajectories but the values of q, j and s do not deviate a lot in different models in the past of the Universe as well as in the near future of the Universe. However, the mean value of the DP parameter gets values higher than 0.5 in the early Universe only in the LVDPt model, which may be indicating that LVDPt can probe the earlier times of the Universe compared to the ΛCDM, DGP, Chaplygin gas and Galileon models. All these models evolve to the de Sitter expansion continuously while the LVDPt model crosses the de Sitter line and reaches the super-exponential expansion rates at some point.

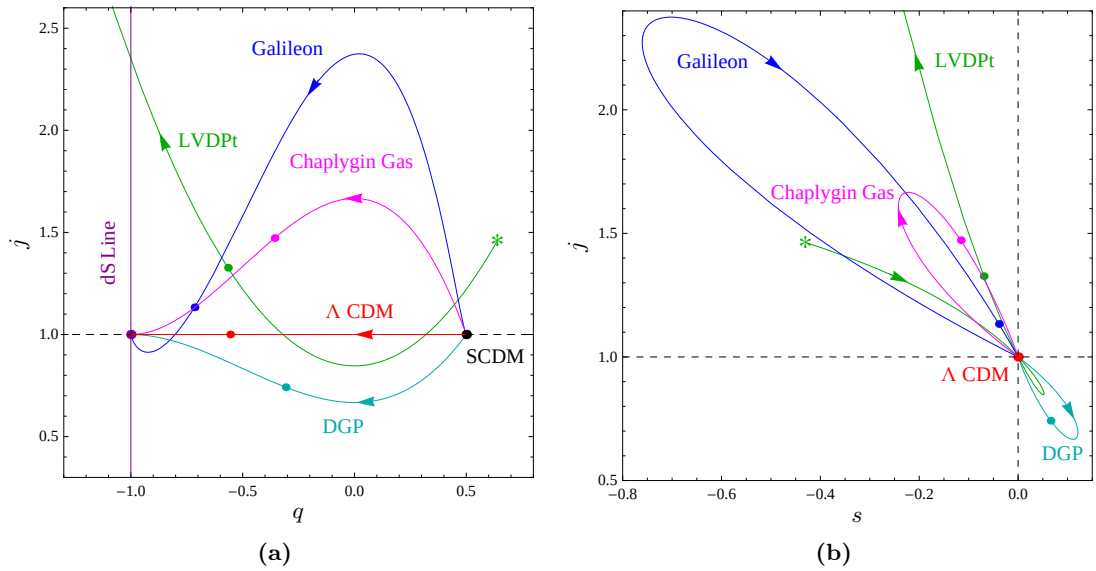


FIG. 2: **(a)** Variation of q versus j. Vertical Purple line stands for the de Sitter (dS) state $q = -1$. **(b)** Variation of s versus j. Horizontal and vertical dashed lines intersect at the ΛCDM point $(0, 1)$. In both panels, the Green curve corresponds to the LVDPt model. The star symbols represent the phase relatively earlier to the matter dominated phase in the LVDPt model. Thereafter, LVDPt $s - j$ curve crosses the ΛCDM statefinder point $(0, 1)$ two times. The final part of the LVDPt $s - j$ curve not shown in panel (b) extends up to the Big Rip point $(-0.943, 11.308)$. Similarly, LVDPt $q - j$ curve in panel (a) crosses the dS line and eventually goes up the Big Rip phase point $(-2.64, 11.308)$. Red, Cyan, Magenta and Blue curves correspond to ΛCDM, DGP, Chaplygin gas and Galileon models respectively. The arrows on the curves show the direction of evolution. The dots on the curves represent the present values of the corresponding (s, j) or (q, j) pair while the black dots show the matter dominated phases of the models.

V. THE HISTORY AND THE FUTURE OF THE UNIVERSE

We discuss some important moments in the expansion history and the possible fates of the Universe in different models; the transition time and redshift from decelerating expansion to the accelerating expansion, the present age of the Universe, the time and redshift when the de Sitter expansion ($q = -1$) will be reached (provided it is reached) and the time when the size of the Universe will become infinite, $a \to \infty$, i.e., when $z \to -1$ (whether in infinite time or in finite time). We give these important moments in the history of the Universe in the LVDPt, ΛCDM, LVDPa and LVDPz models in Table VIII, in the order from left to right according to the goodness of their fit to the observational data. We note that the age of the Universe and the redshift value when the accelerating expansion started and the age of the present Universe are consistent in LVDPt, ΛCDM and LVDPa models, though all these models predict quite different futures for the Universe.

We leave the LVDPa out of our discussion since in this model $q \to -\infty$ as $a \to \infty$, which does not seem realistic. We, however, calculate numerically and give in Table VIII, the time and the redshift when the de Sitter expansion rate is reached and the time when the Universe will reach infinitely large sizes. On the other hand, although LVDPz is the most unreliable law (because of its poor fit to the observational data compared to the others) it is worth checking out the future of the Universe in this model, since it is the only unbiased model for a certain future of the Universe and give us an idea whether the observational data favors finite or infinite future for the Universe. We note that within 1σ error region, the LVDPz model excludes neither the forever expanding Universe ($q_{z=-1} \geq -1$) nor a Universe with a finite lifetime $q_{z=-1} < -1$. Considering the mean value $q_{z=-1} = -1.130 < -1$, on the other hand, we see that it favors a Big Rip end for the Universe. We calculate numerically and give in Table VIII the time and redshift when the de Sitter expansion will be reached as well as when Big Rip will happen considering the mean values of the LVDPz parameters given in Table III. We also plot the $\{q_0, q_1\}$ plane with 68.3 %, 95.4 % and 99.7 % confidence contours, which is divided into two regions by the de Sitter expansion line ($q_0 - q_1 = -1$) in Fig. 3. The $\{q_0, q_1\}$ pairs above the de Sitter line lead to a Big Rip end for the Universe. We are not able to make a discussion on the future of the Universe using the ΛCDM alone since it gives us no option other than a continuous approach to the de Sitter expansion that will never end. Thinking of ΛCDM model fits the observational data better than the LVDPz model (and its shortcomings we discussed previously), we tend to take the future predicted by the ΛCDM model as more reliable. However, we would like to remind that the LVDPt model fits the observational data well in comparison to other models under consideration, and predicts an inevitable Big Rip

TABLE VIII: GoF of LVDPt, ΛCDM, LVDPa and LVDPz models with $H(z)$+SN Ia data and some key moments in the evolution of the Universe in these models. Higher the GoF the better fit the data.

Parameters	LVDPt	ΛCDM	LVDPa	LVDPz
GoF	92.137	92.133	92.129	91.763
Transition time (Gyr)	7.148 ± 0.923	7.233 ± 0.225	7.056 ± 0.754	5.430 ± 0.529
Transition redshift	$0.703^{+0.356}_{-0.148}$	0.682 ± 0.082	$0.662^{+0.198}_{-0.107}$	$0.733^{+0.148}_{-0.095}$
Present age (Gyr)	13.460 ± 0.899	13.389 ± 0.289	13.138 ± 0.754	11.934 ± 0.535
dS time (Gyr)	18.301 ± 1.054	∞	16.344 ± 0.839	40.311 (may be)
dS redshift	$-0.281^{+0.120}_{-0.168}$	-1	$-0.200^{+0.088}_{-0.122}$	-0.799 (may be)
Big Rip time (Gyr)	36.602 ± 7.930	No	27.434 ± 4.015	205.432 (may be)

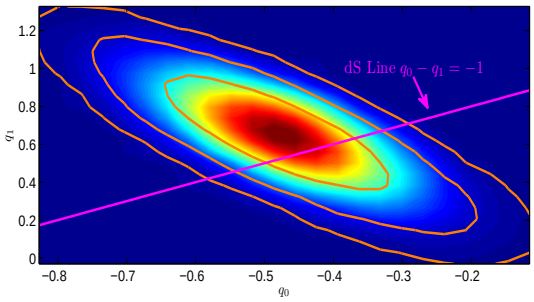


FIG. 3: 2D contours with 68.3 %, 95.4 % and 99.7 % confidence levels are shown for q_0 and q_1 in the LVDPz model.The line $q_0 - q_1 = -1$ across the contours corresponds to the de Sitter phase ($z = -1$) in the LVDPz model. The $\{q_0, q_1\}$ pairs above the de Sitter line lead to Big Rip end of the LVDPz Universe.

end for the Universe. We calculate and give the time and redshift when the de Sitter expansion is reached in LVDPt and when Big Rip will happen in Table VIII. Accordingly, the Universe will start to expand with super-exponential expansion rates ~ 5 billions years after the present time and will reach infinitely large sizes when the Universe will be ~ 36 billions years old.

Considering phantom energy dominated Universe with EoS $w = -3/2$, Caldwell et al. [43] studied history and future of the Universe that ends with a Big Rip. On the other hand, recently Li et al. [21] explored the Big Rip fate of the Universe by using a divergence-free parametrization for the EoS parameter of dark energy. Here, we discuss the future events in the Universe that obeys the LVDPt model assuming general relativity to be the true theory of gravitation. In general relativity, if the Universe is expanding with a DP value less than -1, then there should be a dark energy source with an EoS parameter less than -1 in the Universe whose energy density will grow so that it eventually begins to strip the bound objects in the Universe. According to general relativity, the source for the gravitational potential is the volume integral of $\rho c^2 + 3p$. On the other hand, for a gravitationally bound system with mass M and radius R, the period of a circular orbit around this system at radius R will be $P = 2\pi(R^3/GM)^{1/2}$, where G is the Newton's constant. So, as pointed out in Ref. [43], this system will become unbound roughly when $-\frac{4\pi}{3}(\rho + 3p/c^2)R^3 = M$ that gives $-(1+3w)H^2 = \frac{8\pi^2}{P^2}$. Since the density of dark energy exceeds the energy density of the object and the repulsive gravity of the phantom energy overcomes the forces holding the object together. Accordingly, we calculate and give the events in the future of the Universe considering the best fit values of the LVDPt parameters from the latest $H(z)$ and SN Ia data in the Table IX.

TABLE IX: Future events in the Universe of LVDPt model.

Time	Event
$t_{\rm BR} - 23.1$ Gyr	Today
$t_{\rm BR} - 0.53$ Gyr	Erase Galaxy Clusters
$t_{\rm BR} - 31.53$ Myr	Destroy Milky Way
$t_{\rm BR} - 1.58$ Months	Unbind Solar System
$t_{\rm BR} - 4.30$ Days	Strip Moon
$t_{\rm BR} - 13.32$ Minutes	Earth Explodes
$t_{\rm BR} - 5 \times 10^{-20}$ s	Dissociate Atoms
$t_{\rm BR} = 36.60$ Gyr	Big Rip

From Table IX, one may see that we are 23.1 Gyr away from the Big Rip moment. The dominant dark energy in phantom region ($w < -1$) will start erasing the galaxy clusters around 0.53 Gyr before the end of the Universe. Milky Way will be destroyed around 31.53 Myr while solar system will be disturbed around 1.58 months before the expiry of the Universe. Moon will

be stripped apart around 4.30 days and destruction of earth will take place just before 13.32 minutes of the tragic end of the Universe. Finally the dissociation of atoms would happen in a fraction of the last second of the life of the Universe and thereby leading to the end of the Universe in Big Rip.

VI. CONCLUDING REMARKS

In this study, we have confronted the three theoretically motivated LVDP laws and the standard ΛCDM model with the latest observational data from $H(z)$ compilation (25 data points) and SN Ia Union2.1 compilation (580 data points). We have investigated in detail the kinematics and the fate of the Universe in all the models under consideration. Some important findings of the study are summarized in the following:

(i) The study reveals that the LVDPt law is superior than LVDPz and LVDPa laws in many aspects. The LVDPt is found to be well behaved throughout the evolution of the Universe in comparison with the LVDPz (ill behavior in the past) and LVDPa (ill behavior in the future) laws. In particular, the GoF to the observational data is found to be the best for the LVDPt law.

(ii) From the kinematics and dynamics (assuming general relativity) of the Universe when studied by considering the LVDPt law in comparison with the standard ΛCDM model, it has been found that these two models are observationally indistinguishable for a considerable passage of time, particularly in the vicinity of the present epoch. For instance, in the LVDPt model, the accelerated expansion of the Universe started when the Universe was 7.148 ± 0.923 Gyr old at redshift $0.703^{+0.356}_{-0.148}$ and the age of the present Universe is 13.460 ± 0.899 Gyr, whereas these values are found to be 7.233 ± 0.225 Gyr, 0.682 ± 0.082 and 13.389 ± 0.289 Gyr in the ΛCDM model. Thus, the LVDPt mimics the ΛCDM behavior of the Universe well in the vicinity of the present epoch. In addition, LVDPt finds slightly better fit to the observational data in contrast with the ΛCDM model. However, these two models make slightly different predictions for the earlier times of the Universe but considerably different predictions on the future of the Universe.

(iii) From the the statefinder analysis of LVDPt model along with the popular models such as ΛCDM, DGP, Chaplygin gas and Galileon, we find that the mean value of the DP gets values higher than $\frac{1}{2}$ in the early Universe only in the LVDPt model, which may be indicating that LVDPt is capable of probing the earlier times of the Universe in comparison to the the popular models under consideration, which evolve from SCDM spot ($q = \frac{1}{2}$). Further, all these models evolve to the de Sitter expansion continuously while the LVDPt model crosses the de Sitter line and reaches the super-exponential expansion rates at some point.

(iv) We find that LVDPt Universe has finite life time and Big Rip is its essential feature. Observational constraints on cosmic doomsday carried out in the context of the LVDPt model reveal that the Universe will end in Big Rip when the age of the Universe is 36.602 ± 7.930 Gyr. The de Sitter time in the LVDPt Universe is exactly half the Big Rip time. This feature of the LVDPt Universe seems interesting and merits further investigation to explore its possible role in solving cosmological problems such as the cosmic-coincidence problem.

Nevertheless, we ultimately cannot be sure about the fate of the Universe as long as we do not experience the future. The most reliable but still not the ultimate prediction on the future of the Universe can be the one predicted through a successfully achieved theory of everything (such as string theory) that can pass all the experiments in the past of the Universe successfully. However, we humankind cannot help asking the fate of the Universe and therefore, temptations would continue to discuss the fate of the Universe through the theoretical models that fit the observational data well.

Acknowledgments

Ö.A. and T.D. appreciate the support given by the Turkish Academy of Sciences (TÜBA). Ö.A. also acknowledges the support from Koç University. S.K. acknowledges the financial support from the Department of Science and Technology (DST), India under project No. SR/FTP/PS-102/2011. L. Xu's work is supported in part by NSFC under the Grant No. 11275035 and "the Fundamental Research Funds for the Central Universities" under the Grant No. DUT13LK01.

Appendix A: Constraint methodology using observational Hubble ($H(z)$) data

The differential ages of the galaxies are used for the observational Hubble data via the relation [58–60]

$$H(z) = -\frac{1}{1+z}\frac{dz}{dt}. \tag{A1}$$

We have used 25 OHD data points shown in Table X, recently compiled by Zhang et al. [52].

The mean values of the model parameters are determined by minimizing

$$\chi^2_{OHD}(p_s) = \sum_{i=1}^{25} \frac{[H_{th}(p_s; z_i) - H_{obs}(z_i)]^2}{\sigma^2_{H(z_i)}}, \tag{A2}$$

where p_s denotes the parameters of the model, H_{th} is the theoretical (model based) value for the Hubble parameter, H_{obs} is the observed one, $\sigma_{H(z_i)}$ is the standard error in the observed value, and the summation runs over the 25 observational Hubble data points at redshifts z_i.

TABLE X: $H(z)(\mathrm{km\,s^{-1}\,Mpc^{-1}})$ measurements and their 1σ errors.

z	$H(z)$	$\sigma_{H(z)}$	Reference
0.090	69	12	[61]
0.170	83	8	[61]
0.270	77	14	[61]
0.400	95	17	[61]
0.900	117	23	[61]
1.300	168	17	[61]
1.430	177	18	[61]
1.530	140	14	[61]
1.750	202	40	[61]
0.480	97	62	[62]
0.880	90	40	[62]
0.179	75	4	[63]
0.199	75	5	[63]
0.352	83	14	[63]
0.593	104	13	[63]
0.680	92	8	[63]
0.781	105	12	[63]
0.875	125	17	[63]
1.037	154	20	[63]
0.24	79.69	3.32	[64]
0.43	86.45	3.27	[64]
0.07	69.0	19.6	[52]
0.12	68.6	26.2	[52]
0.20	72.9	29.6	[52]
0.28	88.8	36.6	[52]

Appendix B: Constraint methodology using observational SN Ia data

The observational SN Ia data have played an important role to discover the kinematic state of our Universe [1, 2]. Recently, Suzuki et al. [53] used the framework devised by Kowalski [65] and created a compilation of 580 supernovae, known as Union2.1 compilation. We use this Union2.1 SN Ia data compilation to constrain the model parameters as follows.

The distance modulus $\mu(z)$ is defined as

$$\mu_{th}(z) = 5\log_{10}[\bar{d}_L(z)] + \mu_0, \tag{B1}$$

where $\mu_0 \equiv 42.38 - 5\log_{10} h$ with $H_0 = 100h$ km s^{-1}Mpc^{-1}. Further, $\bar{d}_L(z) = H_0 d_L(z)/c$ is the Hubble-free luminosity, where

$$d_L(z) = \frac{c(1+z)}{H_0\sqrt{|\Omega_k|}}\mathrm{S}\left[\sqrt{|\Omega_k|}\int_0^z \frac{dz'}{E(z')}\right] \tag{B2}$$

with $E^2(z) = H^2(z)/H_0^2$, and $S(x) = \sinh x$ for $k = -1$, $S(x) = x$ for $k = 0$ and $S(x) = \sin x$ for $k = 1$. The observed distance modulus $\mu_{obs}(z_i)$ of SN Ia at redshift z_i is given by

$$\mu_{obs}(z_i) = m_{obs}(z_i) - M, \tag{B3}$$

where M is the absolute magnitude.

The mean values of the model parameters p_s using the SN Ia data set are determined by minimizing the quantity

$$\begin{aligned}\chi^2(p_s, M') &\equiv \sum_{i=1}^{580} \frac{\{\mu_{obs}(z_i) - \mu_{th}(p_s, z_i)\}^2}{\sigma_i^2} \\ &= \sum_{i=1}^{580} \frac{\left\{5\log_{10}[\bar{d}_L(p_s, z_i)] - m_{obs}(z_i) + M'\right\}^2}{\sigma_i^2},\end{aligned} \tag{B4}$$

where $M' \equiv \mu_0 + M$ being a nuisance parameter can be marginalized over analytically [66] as

$$\bar{\chi}^2(p_s) = -2\ln\int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2}\chi^2(p_s, M')\right] dM',$$

resulting to

$$\bar{\chi}^2 = A - \frac{B^2}{C} + \ln\left(\frac{C}{2\pi}\right), \tag{B5}$$

with

$$A = \sum_{i,j}^{SN} \left\{5\log_{10}[\bar{d}_L(p_s, z_i)] - m_{obs}(z_i)\right\} \cdot \mathrm{Cov}_{ij}^{-1} \cdot \left\{5\log_{10}[\bar{d}_L(p_s, z_j)] - m_{obs}(z_j)\right\},$$

$$B = \sum_{i,j}^{SN} \mathrm{Cov}_{ij}^{-1} \cdot \left\{5\log_{10}[\bar{d}_L(p_s, z_j)] - m_{obs}(z_j)\right\},$$

$$C = \sum_{i}^{SN} \mathrm{Cov}_{ii}^{-1}.$$

In the above, Cov_{ij}^{-1} is the inverse of covariance matrix with or without systematic errors [53]. Relation (B4) has a minimum at $M' = B/C$, which depends of the values of h and M. Finally, the expression

$$\chi^2_{SN}(p_s) = A - \frac{B^2}{C}, \qquad \text{(B6)}$$

which matches (B5) up to a constant, is often used in the likelihood analysis [66, 67]. Thus, in this case the results would remain unaffected by a flat M' distribution. It may be noted that the results will be different with or without the systematic errors. In this study, all results are derived considering the systematic errors.

[1] A.G. Riess et al. [Supernova Search Team Collaboration], *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, Astron. J. **116** (1998) 1009-1038

[2] S. Perlmutter et al. [Supernova Cosmology Project Collaboration], *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*, Astrophys. J. **517** (1999) 565-586

[3] E.J. Copeland, M. Sami, S. Tsujikawa, *Dynamics of Dark Energy*, Int. J. Mod. Phys. D **15** (2006) 1753-1936

[4] K. Bamba, S. Capozziello, S. Nojiri, S.D. Odintsov, *Dark energy cosmology: the equivalent description via different theoretical models and cosmography tests*, Astrophys. Space Sci. **342** (2012) 155-228

[5] S. Capozziello, M. De Laurentis, *Extended Theories of Gravity*, Phys. Rep. **509** (2011) 167-321

[6] S. Nojiri and S.D. Odintsov, *Unified cosmic history in modified gravity: From F(R) theory to Lorentz non-invariant models*, Phys. Rep. **505** (2011) 59-144

[7] T. Clifton, P.G. Ferreira, A. Padilla, C. Skordis, *Modified Gravity and Cosmology*, Phys. Rep. **513** (2012) 1-189

[8] Ø. Grøn, S. Hervik, **Einstein's General Theory of Relativity: With Modern Applications in Cosmology**, Springer, New York (2007)

[9] Ya B. Zeldovich, *The Cosmological Constant and the Theory of Elementary Particles*, Sov. Phys. Usp. **11** (1968) 381-393

[10] S. Weinberg, *The cosmological constant problem*, Rev. Mod. Phys. **61** (1989) 1-23

[11] V. Sahni, T.D. Saini, A.A. Starobinsky, *The Case for a Positive Cosmological Λ-Term*, Int. J. Mod. Phys. D **9** (2000) 373-443

[12] P. A. R. Ade et al. [Planck Collaboration], *Planck 2013 results. XVI. Cosmological parameters*, arXiv:1303.5076 [astro-ph.CO]

[13] A.G. Riess, L. Macri, S. Casertano et al., *A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3*, Astrophys. J. **730** (2011) 119

[14] M.S. Turner, A.G. Riess, *Do Type Ia Supernovae Provide Direct Evidence for Past Deceleration of the Universe?*, Astrophys. J. **569** (2002) 18-22

[15] D. Rapetti, S.W. Allen, M.A. Amin, R.D. Blandford, *A kinematical approach to dark energy studies*, Mon. Not. R. Astron. Soc. **375** (2007) 1510-1520

[16] E.E.O. Ishida, R.R.R. Reis, A.V. Toribio, I. Waga, *When did cosmic acceleration start? How fast was the transition?*, Astropart. Phys. **28** (2008) 547-552

[17] E. Mörtsell, C. Clarkson, *Model independent constraints on the cosmological expansion rate*, J. Cosmol. Astropart. Phys. **01** (200) 044

[18] L. Xu, W. Li, J. Lu, *Constraints on kinematic model from recent cosmic observations: SN Ia, BAO and observational Hubble data*, J. Cosmol. Astropart. Phys. **07** (2009) 031

[19] J. Lu, L. Xu, M. Liu, *Constraints on kinematic models from the latest observational data*, Phys. Lett. B **699** (2011) 246-250

[20] E.M. Barboza, F.C. Carvalho, *A kinematic method to probe cosmic acceleration*, Phys. Lett. B **715** (2012) 19-24

[21] X.D. Li, S. Wang, Q.G. Huang, X. Zhang, M. Li, *Dark energy and fate of the Universe*, Sci. China Phys. Mech. Astron. **55** (2012) 1330

[22] O. Farooq, B. Ratra, *Hubble Parameter Measurement Constraints on the Cosmological Deceleration-Acceleration Transition Redshift*, Astrophys. J. **766** (2013) L7

[23] A.G. Riess et al., *Type Ia Supernova Discoveries at $z > 1$ from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution*, Astrophys. J. **607** (2004) 665-687

[24] J.V. Cunha, J.A.S. Lima, *Transition redshift: new kinematic constraints from supernovae*, Mon. Not. R. Astron. Soc. **390** (2008) 210-217

[25] L. Xu, H. Liu, *Constraints to DPs by Recent Cosmic Observations*, Mod. Phys. Lett. A **23** (2008) 1939-1948

[26] L. Xu, J. Lu, *Cosmic Constraints on deceleration parameter with Sne Ia and CMB*, Mod. Phys. Lett. A **24** (2009) 369-376

[27] J.V. Cunha, *Kinematic constraints to the transition redshift from supernovae type Ia union data*, Phys. Rev. D **79** (2009) 047301

[28] R. Nair, S. Jhingan, D. Jain, *Cosmokinetics: A joint analysis of Standard Candles, Rulers and Cosmic Clocks*, J. Cosmol. Astropart. Phys. **01** (2012) 018

[29] S. del Campo, I. Duran, R. Herrera, D. Pavón, *Three thermodynamically based parametrizations of the DP*, Phys. Rev. D **86** (2012) 083509

[30] Ö. Akarsu, T. Dereli, *Cosmological Models with Linearly Varying DP*, Int. J. Theor. Phys. **51** (2012) 612-621

[31] Ö. Akarsu, T. Dereli, *A Comparison of the LVDP and ΛCDM Cosmological Models*, Int. J. Theor. Phys. **51** (2012) 2995-3000

[32] S. Kumar, *Observational constraints on Hubble constant and DP in power-law cosmology*, Mon. Not. R. Astron. Soc. **422** (2012) 2532-2538

[33] Y. Chen, B. Ratra, *Hubble parameter data constraints on dark energy*, Phys. Lett. B **703** (2011) 406-411

[34] R. Amanullah et al., *Spectra and Hubble Space Telescope Light Curves of Six Type Ia Supernovae at 0.511 < z < 1.12 and the Union2 Compilation*, Astrophys. J. **716** (2010) 712-738

[35] V. Sahni, T.D. Saini, A.A. Starobinsky, U. Alam, *Statefinder - A new geometrical diagnostic of dark energy*, JETP Letters **77** (2003) 201-206

[36] M. Visser, *Jerk, snap and the cosmological equation of state*, Classical Quant. Grav. **21** (2004) 2603-2615

[37] C. Cattoen, M. Visser, *Cosmographic Hubble fits to the supernova data*, Phys. Rev. D **78** (2008) 063501

[38] F.Y. Wang, Z.G. Dai, S. Qi, *Probing the cosmographic parameters to distinguish between dark energy and modified gravity models*, Astron. Astrophys. **507** (2009) 53-59

[39] V. Vincenzo et al., *High-redshift cosmography*, J. Cosmol. Astropart. Phys. **03** (2010) 005

[40] S. Capozziello, R. Lazkoz, V. Salzano, *Comprehensive cosmographic analysis by Markov chain method*, Phys. Rev. D **84** (2011) 124061

[41] J.Q. Xia, V. Vitagliano, S. Liberati, M. Viel, *Cosmography beyond Standard Candles and Rulers*, Phys. Rev. D **85** (2012) 043520

[42] V. Simha, G. Steigman, *Constraining The Early-Universe Baryon Density And Expansion Rate*, J. Cosmol. Astropart. Phys. **06** (2008) 016

[43] R.R. Caldwell, M. Kamionkowski, N.N. Weinberg, *Phantom Energy: Dark Energy with $w < -1$ Causes a Cosmic Doomsday*, Phys. Rev. Lett. **91** (2003) 071301

[44] S. Nesseris, L. Perivolaropoulos, *Crossing the phantom divide: theoretical implications and observational status*, J. Cosmol. Astropart. Phys. **0701** (2007) 018

[45] Y.F. Cai, E.N. Saridakis, M.R. Setare, J.Q. Xia, *Quintom cosmology: Theoretical implications and observations*, Phys. Rep. **493** (2010) 1-60

[46] W. Godlowski, M. Szydlowski, *How many parameters in the cosmological models with dark energy?*, Phys. Lett. B **623** (2005) 10-16.

[47] G.B. Zhao et al., *Probing for dynamics of dark energy and curvature of Universe with latest cosmological observations*, Phys. Lett. B **648** (2007) 8-13

[48] J.A. Vazquez, M. Bridges, M.P. Hobson, A.N. Lasenby, *Reconstruction of the dark energy equation of state*, J. Cosmol. Astropart. Phys. **09** (2012) 020

[49] B. Novosyadlyj, O. Sergijenko, R. Durrer, V. Pelykh, *Do the cosmological observational data prefer phantom dark energy?*, (2012) arXiv:1206.5194 [astro-ph.CO]

[50] D. Parkinson et al., *The WiggleZ Dark Energy Survey: Final data release and cosmological results*, (2012) arXiv:1210.2130 [astro-ph.CO]

[51] Ö. Akarsu, T. Dereli, *Late time acceleration of the 3-space in a higher dimensional steady state Universe in dilaton gravity*, J. Cosmol. Astropart. Phys. **02** (2013) 50

[52] C. Zhang, et al., *Four New Observational $H(z)$ Data From Luminous Red Galaxies Sloan Digital Sky Survey Data Release Seven*, (2012) arXiv:1207.4541 [astro-ph.CO]

[53] N. Suzuki, et al., *The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above $z > 1$ and Building an Early-type-hosted Supernova Sample*, Astropart. Phys. **746** (2012) 85

[54] A. Lewis, S. Bridle, *Cosmological parameters from CMB and other data: A Monte Carlo approach*, Phys. Rev. D **66** (2002) 103511 (http://cosmologist.info/cosmomc/)

[55] C.L. Bennett et al., *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, (2012) arXiv:1212.5225 [astro-ph.CO]

[56] V. Sahni, T.D. Saini, A.A. Starobinsky, U. Alam, *Statefinder-A new geometrical diagnostic of dark energy*, JETP Lett. **77** (2003) 201-206

[57] M. Sami, M. Shahalam, M. Skugoreva, A. Toporensky, *Cosmological dynamics of a nonminimally coupled scalar field system and its late time cosmic relevance*, Phys. Rev. D **86** (2012) 103532

[58] R. Jimenez and A. Loeb, *Constraining Cosmological Parameters Based on Relative Galaxy Ages*, Astrophys. J. **573** (2002) 37-42

[59] R. Jimenez, L. Verde, T. Treu, D. Stern, *Constraints on the equation of sate of Dark Energy and the Hubble Constant from Stellar Ages and the Cosmic Microwave Background*, Astrophys. J. **593** (2003) 622-629

[60] Z. Yi, T. Zhang, *Constraints on Holographic Dark Energy Models Using the Differential Ages of Passively Evolving Galaxies*, Mod. Phys. Lett. A **22** (2007) 41-53

[61] J. Simon, L. Verde, R. Jimenez, *Constraints on the redshift dependence of the dark energy potential*, Phys. Rev. D **71** (2005) 123001

[62] D. Stern, R. Jimenez, L. Verde, M. Kamionkowski, S. A. Standford, *Cosmic chronometers: constraining the equation of state of dark energy. I: H(z) measurements*, J. Cosmol. Astropart. Phys. **02** (2010) 008

[63] M. Moresco, A. Cimatti, R. Jimenez et al., *Improved constraints on the expansion rate of the Universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers*, J. Cosmol. Astropart. Phys. **08** (2012) 006

[64] E. Gaztanaga, A. Cabre, L. Hui, *Clustering of luminous red galaxies - IV. Baryon acoustic peak in the line-of-sight direction and a direct measurement of H(z)*, Mon. Not. R. Astron. Soc. **399** (2009) 1663-1680

[65] M. Kowalski et al., *Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets*,

Astropart. Phys. **686** (2008) 749

[66] S. Nesseris, L. Perivolaropoulos, *Comparison of the legacy and gold type Ia supernovae dataset constraints on dark energy models*, Phys. Rev. D **72** (2005) 123519; L. Perivolaropoulos, *Constraints on linear negative potentials in quintessence and phantom models from recent supernova data*, Phys. Rev. D **71** (2005) 063503; E. Di Pietro, J. F. Claeskens, *Future supernovae data and quintessence models*, Mon. Not. Roy. Astron. Soc. **341** (2003) 1299; A.C.C. Guimaraes, J.V. Cunha, J.A.S. Lima, *Bayesian analysis and constraints on kinematic models from union SNIa*, J. Cosmol. Astropart. Phys. **10** (2009) 010

[67] E. Garcia-Berro, E. Gaztanaga, J. Isern, O. Benvenuto, L. Althaus, *On the Evolution of Cosmological Type Ia Supernovae and the Gravitational Constant*, (1999) arXiv:astro-ph/9907440; A. Riazuelo, J. Uzan, *Cosmological observations in scalar-tensor quintessence*, Phys. Rev. D **66** (2002) 023525; V. Acquaviva, L. Verde, *Observational signatures of Jordan Brans Dicke theories of gravity*, J. Cosmol. Astropart. Phys. **12** (2007) 001

Journal of Statistical Theory and Applications, Vol. 12, No. 1 (May 2013), 120-128

On Discrete Residual and Past Measures of Inaccuracy

Vikas Kumar, Richa Thapliyal, and H.C. Taneja

Department of Applied Mathematics,
Delhi Technological University,
Bawana Road, Delhi-110042, India

vikas_iitr82@yahoo.co.in, richa31aug@gmail.com, hctaneja@rediffmail.com

The present communication considers a discrete dynamic measure of inaccuracy between two residual and past lifetime distributions. Under the assumption that the true distribution F and reference distribution G satisfy the proportional hazard model (PHM) and proportional reversed hazard model (PRHM), it has been shown that the proposed measures determines the lifetime distribution uniquely. A relation between Kerridge inaccuracy, residual inaccuracy and past inaccuracy has been explored.

Keywords: Kerridge inaccuracy, dynamic inaccuracy, residual entropy, residual lifetime, proportional hazard model.

1. Introduction

Let X and Y be two non-negative random variables representing time to failure of two systems with p.d.f. respectively $f(x)$ and $g(x)$. Let $F(x) = P(X \leqslant x)$ and $G(Y) = P(Y \leqslant y)$ be failure distributions, and $\overline{F}(x) = 1 - F(x)$, $\overline{G}(x) = 1 - G(x)$ be survival functions of X and Y respectively. Shannon's (1948) measure of uncertainty associated with the random variable X and Kerridge measure of inaccuracy (1961) are given as

$$H(f) = -\int_0^\infty f(x) \log f(x) dx, \tag{1.1}$$

and

$$H(f;g) = -\int_0^\infty f(x) \log g(x) dx, \tag{1.2}$$

respectively. In case $g(x) = f(x)$, then (1.2) reduces to (1.1).

The measures (1.1) and (1.2) are not applicable to a system which has survived for some unit of time. Ebrahimi (1996) considered the entropy of the residual lifetime $X_t = [X - t \mid X > t]$ as a

residual measure of uncertainty given by

$$H(f;t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} dx. \tag{1.3}$$

Extending the information measure (1.3), a residual measure of inaccuracy, refer to Taneja *et al.* (2009), is given as

$$H(f,g;t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{g(x)}{\overline{G}(t)} dx. \tag{1.4}$$

Here $g(x)$ can be interpreted as some reference p.d.f., whereas $f(x)$ is the baseline p.d.f. in context with a system under consideration. Further, the residual measure of inaccuracy given by (1.4) has been extended to weighted (length biased) residual inaccuracy measure, refer to Kumar *et al.* (2010).

There are many situations where a continuous time is inappropriate for describing the lifetime of devices and other systems, refer to Xekalaki (1983). Some specific situations worth mentioning are: (i) actuaries and bio-statisticians are interested in the lifetimes of persons or organisms, measured in months, weeks, or days, (ii) in case of equipment operating in cycles the random variable of interest is the successful number of cycles before the failure. For instance, the number of flashes in a car-flasher prior to failure of the device, (iii) a discrete life distribution is a natural choice where failure occurs only due to incoming shocks, for example, in weapons reliability, the numbers of rounds fired until failure is more important than age at failure. Further discrete lifetime also occurs through grouping or finite measurement of continuous time phenomena. Since there is a limit on the precision of any measurement, it can be arguably said that sample from a continuous distribution exist only in theory.

Let X be a discrete lifetime random variable taking values on $N = \{1,2,3,\ldots,n\}$, $1 \leqslant n < \infty$, with probability $P(X = i) = p_i$, $0 \leqslant p_i \leqslant 1$, $\sum_{i=1}^n p_i = 1$. Let $F(i) = P(X \leqslant i) = \sum_{j=0}^{j=i} p_j$, be the lifetime distribution function and $\overline{F}(i) = 1 - F(i-1)$, $i = 1,2,\ldots$ the corresponding survival function. Shannon (1948) introduced the measure of information (or, uncertainty) associated with X given as

$$H(P) = -\sum_{i=1}^n p_i \log p_i. \tag{1.5}$$

Next, consider a situation where an experimenter needs to predict the probabilities of various outcomes in an experiment. Let Y be the reference (or, predicted) random variable taking values on $N = \{1,2,3,\ldots,n\}$, $1 \leqslant n < \infty$ with probability $P\{Y = i\} = q_i$, $0 \leqslant q_i \leqslant 1$, $\sum_{i=1}^n q_i = 1$, while the actual distribution is given by $P\{X = i\} = p_i$, that is the experimenter assumes that the probability of the i^{th} outcome is q_i whereas the true probability is p_i. Then the average inaccuracy in the statement of the observer, as proposed by Kerridge (1961), can be measured by

$$H(P,Q) = -\sum_{i=1}^n p_i \log q_i. \tag{1.6}$$

The modification of Shannon's entropy as a measure of uncertainty in residual lifetime distribution for discrete random variable has drawn attention of many researchers. Rajesh and Nair (1998) have

proposed

$$H(P;j) = -\sum_{k=j}^{\infty} \frac{p(k)}{\overline{F}(j)} \log \frac{p(k)}{\overline{F}(j)}, \tag{1.7}$$

as residual lifetime information measure in discrete domain and showed that $H(P;j)$ uniquely determines the distribution function $F(t)$ and characterized the same. Generalizing the concept of Rajesh and Nair (1998), Belzuence (2004) extended γ-entropy, refer to Ord *et al.* (1981), for the residual lifetime in discrete domain, and showed that $H_\gamma(P;j)$ uniquely determines the lifetime distribution. Nanda and Paul (2006) have extended this to entropy of order β and have studied their properties. In this communication we propose and study the measures of dynamic (residual and past both) inaccuracy, when the distributions P and Q are discrete domain, and satisfy *proportional hazard model* (PHM), refer to Cox (1959), or *proportional reversed hazard model* (PRHM), refer to Gupta and Gupta (2007). In Section 2, we consider the concept of PHM and PRHM in case of discrete domain. In Section 3, we propose and characterize the residual inaccuracy measure under proportional hazard model (PHM). In Section 4, we propose and characterize the past inaccuracy measure under proportional reversed hazard model (PRHM).

2. Discrete Case Hazard Models

The hazard rate and the reverse hazard rate functions for a random variable X with non-negative integral support are defined by

$$h(i) = \frac{p(i)}{\overline{F}(i)}, \quad i = 1,2,\ldots \tag{2.1}$$

and

$$r(i) = \frac{p(i)}{F(i)}, \quad i = 1,2,\ldots \tag{2.2}$$

respectively. In the discrete setup both the hazard rate and the reversed hazard rate can be interpreted as a probability which is not the case in the continuous case. If X represents the lifetime of a component then $h(i)$ is the probability that the component will fail at time $X = i$ given that it has survived up to the time before i. Similarly, $r(i)$ is the probability that the component will fail at time $X = i$, given it is known to have failed before i. But many of the properties of the hazard rate and reversed hazard rate which hold in the continuous case do not hold in the discrete case. Specifically the definitions (2.1) and (2.2) don't lead respectively to PHM and PRHM in discrete domain. Xie, Gaudoin and Bracquemond (2002) have redefined the hazard rate and reversed hazard rate functions in discrete setup as follows:

$$\lambda(i) = \log \frac{\overline{F}(i-1)}{\overline{F}(i)}, \quad i = 1,2,\ldots \tag{2.3}$$

and

$$\mu(i) = \log \frac{F(i)}{F(i-1)}, \quad i = 1,2,\ldots \tag{2.4}$$

It is easy to verify that the functions defined by (2.1) and (2.3) satisfy

$$h(i) = 1 - \exp[-\lambda(i)], \tag{2.5}$$

and defined by (2.2) and (2.4) lead to

$$r(i) = 1 - \exp[-\mu(i)], \ \ \forall\, i = 1,2,\ldots \tag{2.6}$$

Further we note that the both hazard rates (as well as reversed hazard rates) preserve the same monotonicity property.

It is easy to verify that if X and Y are two discrete random variables with non-negative integral support with probability distribution functions F and G respectively, then the hazard rate function defined by (2.3) satisfy the PHM

$$\overline{G}(i) = \left\{\overline{F}(i)\right\}^{\beta}, \ \ \beta > 0, \tag{2.7}$$

equivalently

$$\lambda^*(i) = \beta\,\lambda(i); \tag{2.8}$$

and the reversed hazard rate function defined by (2.4) satisfy the PRHM

$$G(i) = \{F(i)\}^{\beta}, \ \ \beta > 0 \tag{2.9}$$

equivalently

$$\mu^*(i) = \beta\,\mu(i). \tag{2.10}$$

3. Discrete Residual Inaccuracy Measure and Characterization Problem

In sequel to measure of inaccuracy (1.6), we propose

$$H(P,Q;j) = -\sum_{k=J}^{\infty} \frac{p(k)}{\overline{F}(j)} \log \frac{q(k)}{\overline{G}(j)}, \tag{3.1}$$

as a discrete residual inaccuracy associated with two residual lifetime distributions P and Q. When $j = 0$, then (3.1) becomes (1.6), and when $p(k) = q(k)$, $\forall\, k$ then (3.1) becomes (1.7), the discrete measure of uncertainty.

Example 3.1. Consider a random variable X with probability mass function (pmf)

$$P(X = i) = p_i = \frac{2i}{n(n+1)}, \quad i = 1,2,\ldots,n,$$

and another random variable Y with pmf

$$P(Y = i) = q_i = \frac{1}{n}, \quad i = 1,2,\ldots,n.$$

Then we have survival function corresponding to X and Y given as

$$\overline{F}(i) = 1 - F(i-1) = 1 - \frac{i(i-1)}{n(n+1)},$$
$$\overline{G}(i) = 1 - G(i-1) = \frac{n-i+1}{n},$$

respectively. Substituting these in (3.1) and simplifying, we obtain the discrete residual inaccuracy measure

$$H(P,Q;j) = \log(n-j+1)\left\{\frac{(n+j)(n-j+1)}{n(n+1)-j(j-1)}\right\}. \tag{3.2}$$

Theorem 3.1. *Let X and Y be two discrete random variables with non-negative integral support satisfying the proportional hazard model* (2.7)*. If $H(P,Q;j)$ is increasing and finite, then $H(P,Q;j)$ uniquely determines the survival function $\overline{F}(j)$ of the random variable X.*

Proof. We have

$$H(P,Q;j) = -\sum_{k=J}^{\infty}\frac{p(k)}{\overline{F}(j)}\log\frac{q(k)}{\overline{G}(j)},$$

which is equivalent to

$$\sum_{k=J}^{\infty}p(k)\log q(k) = \overline{F}(j)\log\overline{G}(j) - \overline{F}(j)H(P,Q;j). \tag{3.3}$$

For $j+1$ we obtain

$$\sum_{k=J+1}^{\infty}p(k)\log q(k) = \overline{F}(j+1)\log\overline{G}(j+1) - \overline{F}(j+1)H(P,Q;j+1). \tag{3.4}$$

Denoting $H(P,Q;j)$ by $H(j)$, and subtracting (3.3)-(3.4), we get

$$p(j)\log q(j) = \overline{F}(j)\left\{\log\overline{G}(j) - H(j)\right\} - \overline{F}(j+1)\left\{\log\overline{G}(j+1) - H(j+1)\right\}.$$

Substituting $p(j) = \overline{F}(j) - \overline{F}(j+1)$, and $q(j) = \overline{G}(j) - \overline{G}(j+1)$, we obtain

$$\begin{aligned}&\left\{\overline{F}(j) - \overline{F}(j+1)\right\}\log\left\{\overline{G}(j) - \overline{G}(j+1)\right\}\\ &\quad= \overline{F}(j)[\log\overline{G}(j) - H(j)] - \overline{F}(j+1)\left\{\log\overline{G}(j+1) - H(j+1)\right\}.\end{aligned} \tag{3.5}$$

Dividing by $\overline{F}(j)$ both side and simplifying we get

$$\begin{aligned}&\left\{\frac{\overline{F}(j+1)}{\overline{F}(j)}\right\}\log\frac{\overline{G}(j+1)}{\overline{G}(j)} + \left\{1 - \frac{\overline{F}(j+1)}{\overline{F}(j)}\right\}\log\left\{1 - \frac{\overline{G}(j+1)}{\overline{G}(j)}\right\}\\ &\qquad= -H(j) - \frac{\overline{F}(j+1)}{\overline{F}(j)}H(j+1).\end{aligned} \tag{3.6}$$

If $e^{-\lambda_j} = \frac{\overline{F}(j+1)}{\overline{F}(j)}$, $e^{-\lambda_j^*} = \frac{\overline{G}(j+1)}{\overline{G}(j)}$, then (3.6) becomes

$$-\lambda_j^* e^{-\lambda_j} + (1-e^{-\lambda_j})\log(1-e^{-\lambda_j^*}) = -H(j) + e^{-\lambda_j}H(j+1). \tag{3.7}$$

Using the proportional hazard model (2.8), (3.7) can be written as

$$(1-e^{-\lambda_j})\log(1-e^{-\beta\lambda_j}) + H(j) - \beta\lambda_j e^{-\lambda_j} - e^{-\lambda_j}H(j+1) = 0. \tag{3.8}$$

It is noted that

$$g(x) = (1-x)\log(1-x^{\beta}) + H(j) + \beta x\log x - xH(j+1) = 0 \tag{3.9}$$

has one root $x = e^{-\lambda_j}$ for fixed j. In this case $g(0) = H(j) \geqslant 0$, $g(1) = H(j) - H(j+1) \leqslant 0$, and

$$g'(x) = -H(j+1) + \log \frac{x^\beta}{1-x^\beta} - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta\,.$$

Therefore, $g(x)$ first decreases and then increases in $(0,1)$, with a minimum at

$$x_j = \left\{ \frac{1}{1+\exp[-H(j+1) - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta]} \right\}^{\frac{1}{\beta}},$$

which implies that equation (3.8) has a unique positive solution $e^{-\lambda_j}$ in $(0,1)$ for all j. Finally, since $1 - e^{-\lambda_j} = h(j)$, where $h(j) = \frac{p(j)}{\overline{F}(j)}$ is the discrete failure rate, thus $H(P,Q;j)$ uniquely determines the distribution function F. □

4. Discrete Past Inaccuracy Measure and Characterization Problem

In many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. In this situation the random variable of interest is ${}_iX = [i - X \mid X \leqslant i]$, known as the *inactivity time*. This is because, once at time X the system fails, and at time i it is observed to be in a failure state, the random time for which the system was down is ${}_iX$, where X is a discrete random variable. Based on this idea Di Crescenzo and Longobardi (2002) have studied the past entropy given by

$$H^*(P;j) = -\sum_{k=0}^{j} \frac{p(k)}{F(j)} \log \frac{p(k)}{F(j)}\,. \tag{4.1}$$

Here we propose a discrete dynamic measure of inaccuracy based on the past entropy over $(0,j)$, defined as

$$H^*(P,Q;j) = -\sum_{k=0}^{j} \frac{p(k)}{F(j)} \log \frac{q(k)}{G(j)}\,, \tag{4.2}$$

When $p(k) = q(k), \forall\, k$ then (4.2) becomes (4.1), the discrete dynamic measure of past entropy given by Di Crescenzo and Longobardi (2002).

Example 4.1. Reconsider the random variable X and Y as defined in Example 3.1, the distribution functions corresponding to X and Y are given by

$$F(j) = \sum_{i=1}^{j} p_i = \frac{j(j+1)}{n(n+1)}\,,$$

and

$$G(j) = \sum_{i=1}^{j} q_i = \frac{j}{n}\,,$$

respectively. Substituting these in (4.2) and simplifying, we obtain the discrete past inaccuracy measure

$$H^*(P,Q;j) = \log j\,. \tag{4.3}$$

Next, we study characterization problem for the discrete past inaccuracy measure under the proportional reversed hazard model. Sengupta *et al.* (1999) illustrated that PRHM leads to a better fit for some data sets than PHM.

Theorem 4.1. *Let X and Y be two non negative random variables with integral support satisfying the proportional reversed hazard model (PRHM) (2.10). Let $H^*(P,Q;j) < \infty$, be an increasing function of j, then $H^*(P,Q;j)$ uniquely determines the survival function $\overline{F}(j)$ of the variable X.*

Proof. From (4.2) we have

$$\sum_{k=0}^{j} p(k)\log q(k) = F(j)\log G(j) - F(j)H^*(P,Q;j). \tag{4.4}$$

By replacing j by $j+1$ in (4.4), we have

$$\sum_{k=0}^{j+1} p(k)\log q(k) = F(j+1)\log G(j+1) - F(j+1)H^*(P,Q;j+1). \tag{4.5}$$

Denoting $H^*(P,Q;j)$ by $H^*(j)$, and subtracting (4.5) from (4.4), we get

$$p(j+1)\log q(j+1) = F(j+1)\left\{\log G(j+1) - H^*(j+1)\right\} - F(j)\left\{\log G(j) - H^*(j)\right\},$$

Writing $p(j+1) = F(j+1) - F(j)$ and $q(j+1) = G(j+1) - G(j)$, we obtain

$$\begin{aligned}&\{F(j+1) - F(j)\}\log\{G(j+1) - G(j)\}\\ &\quad = F(j+1)[\log G(j+1) - H^*(j+1)] - F(j)\{\log G(j) - H^*(j)\}.\end{aligned} \tag{4.6}$$

which can be expressed as

$$(1-\mathrm{e}^{-\mu_j})\log(1-\mathrm{e}^{-\mu_j^*}) - \mu_j^*\mathrm{e}^{-\mu_j} = \mathrm{e}^{-\mu_j}H^*(j) - H^*(j+1), \tag{4.7}$$

where $\mathrm{e}^{-\mu_j} = \frac{F(j)}{F(j+1)}$ and $\mathrm{e}^{-\mu_j^*} = \frac{G(j)}{G(j+1)}$.
Using the proportional reversed hazard model (2.10), then (4.7) becomes

$$(1-\mathrm{e}^{-\mu_j})\log(1-\mathrm{e}^{-\beta\mu_j}) + H^*(j+1) - \beta\mu_j\mathrm{e}^{-\mu_j} - \mathrm{e}^{-\mu_j}H^*(j) = 0. \tag{4.8}$$

It is noted that

$$y_1(x) = (1-x)\log(1-x^\beta) + \beta x\log x - xH^*(j) + H^*(j+1) = 0, \tag{4.9}$$

has one root $x = \mathrm{e}^{-\mu_j}$, for fixed j.
In this case $y_1(0) = H^*(j+1) \geqslant 0$, $y_1(1) = H^*(j+1) - H^*(j) \leqslant 0$ and

$$y_1'(x) = -H^*(j) + \log\frac{x^\beta}{1-x^\beta} - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta\,.$$

Therefore, $y_1(x)$ first decreases and then increases in $(0,1)$, with a minimum at

$$x_j = \left\{\frac{1}{1+\exp[-H^*(j) - \frac{\beta x^{\beta-1}(1-x)}{(1-x^\beta)} + \beta]}\right\}^{\frac{1}{\beta}},$$

which implies that equation (4.9) has a unique positive solution μ_j in $(0,1)$ for all j. Finally, since $1-e^{-\mu_j} = r(j+1)$, where $r(j) = \frac{p(j)}{\overline{F}(j)}$ is the discrete reversed hazard rate, thus $H^*(P,Q;j)$ uniquely determines the distribution function F. This complete the proof. □

Remark 4.1. 1. We observe the following relation between the three inaccuracy measures (1.6), (3.1) and (4.2) as considered above.

$$H(P;Q) = \overline{F}(j+1)H(j+1) + F(j)H^*(j) + H[F(j),G(j)], \tag{4.10}$$

where $H[F(j),G(j)] = -F(j)\log G(j) - [1-F(j)]\log[1-G(j)]$, corresponds to the Kerridge inaccuracy (1961), and $H(p,1-p) = -p\log p - (1-p)\log(1-p)$ is the entropy of a Bernoulli distribution.

2. For $\beta = 1$, value of x_j for residual and past inaccuracy reduces to

$$x_j = \left\{ \frac{1}{1+e^{-}H(j+1)} \right\},$$

and

$$x_j = \left\{ \frac{1}{1+e^{-}H^*(j)} \right\},$$

respectively, these results given by Belzunce *et al.* (2004) and Nanda and Paul (2006) in context with residual and past entropy in discrete domain.

5. Conclusion

So far, discrete life time distributions play only a marginal role in reliability analysis. However, real world life time is either observed in discrete time points or it is measured by discrete quantities. Therefore, the focus of reliability analysis should turn to the more realistic discrete life time distributions with finite support that so far are hardly investigated in academic reliability theory, despite the fact that in real world everything is finite.

The proportional hazards and reversed hazards model are very common in reliability analysis. Here we using these models relevant to situations when lifetimes are discrete and study the characterization problem of discrete dynamic inaccuracy measures.

References

[1] A. Di Crescenzo, M. Longobardi, Entropy-based measure of uncertainty in past lifetime distributions, *Journal of Applied Probability*, **39** (2002) 434–440.
[2] A.K. Nanda, P. Paul, Some properties of past entropy and their applications, *Metrika*, **64** (2006) 47–61.
[3] A.K. Nanda, P. Paul, Some results on generalized residual entropy, *Information Sciences*, **176** (2006) 27–47.
[4] B. Abraham, P.G. Sankaran, Renyi's entropy for residual life distribution, *Statistical Papers*, **46** (2005) 17–30.
[5] C.E. Shannon, A mathematical theory of communications, *Bell System Technical Journal*, **27** (1948) 379–423.
[6] D.F. Kerridge, Inaccuracy and inference, *Journal of Royal Statistical Society, Series B*, **23** (1961) 184–194.

[7] D.R. Cox, The analysis of exponentially distributed lifetimes with two type of failure, *Journal of Royal Statistical Society, Series B*, **21** (1959) 411–421.
[8] D. Sengupta, H. Singh, A.K. Nanda, *The proportional reversed hazards model*, Technical Report, Applied Statistics Division, Indian statistical Institute, Kolkata (1999).
[9] E. Xekalaki, Hazard function and life distributions in discrete time, *Communications in Statistics: Theory and Methods*, **12** (1983) 2503–2509.
[10] F. Belzunce, J. Navarro, J.M. Ruiz, Y. Aguila, Some results on residual entropy function, *Metrika*, **59** (2004) 147–161.
[11] G. Rajesh, K.R.M. Nair, Residual entropy function in discrete time, *Far East Journal of Theoretical Statistics*, **2** (1998) 43–57.
[12] H.C. Taneja, V. Kumar, R. Srivastava, A dynamic measure of inaccuracy between two residual lifetime distributions, *International Mathematical Forum*, **4** (25) (2009) 1213–1220.
[13] J.K. Ord, G.P. Patil, C. Taillie, The choice of a distribution to describe personal incomes. In: Taillie C., Patil G.P., Baldessari, B.A. (eds.) *Statistical Distributions in Scientific Work*, Vol. 6, D. Reidel Publishing Co., Dordrecht-Boston, Massachusetts, (1981) pp. 193–201.
[14] M. Xie, O. Gaudoin, C. Bracquemond, Redefining failure rate function for discrete distributions, *International Journal of Reliability Quality Safety Engineering*, **9** (2002) 275–285.
[15] N. Ebrahimi, How to measure uncertainty about residual lifetime, *Sankhya A*, **58** (2002) 48–57.
[16] R.C. Gupta, R.D. Gupta, Proportional reversed hazard rate model and its applications, *Journal of Statistical Planning and Inference*, **137** (2007) 3525–3536.
[17] V. Kumar, H.C. Taneja, R. Srivastava Length biased weighted residual inaccuracy measure, *Metron*, **LXIII 2** (2010) 153–160.

IOSR Journal of Applied Physics (IOSR-JAP)
e-ISSN: 2278-4861. Volume 3, Issue 5 (Mar. - Apr. 2013), PP 67-79
www.iosrjournals.org

Silicon Photonics: A Review

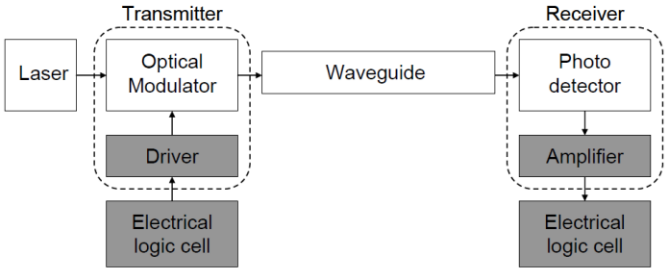
Ashish Dhiman
[1] (*Department of Electronics and Communication, Delhi Technological University (Formerly Delhi College of Engineering), Delhi, India)*

Abstract *: Previously, researchers were trying to replace the copper wires with the optical technology. They used photons instead of electrons to carry out the data within machines. The future of IC industry depends upon the speed and bandwidth requirements, but in metallic interconnections we have constraints on speed as well as bandwidth. With the advancement in computer technology we are now more dependent upon ultra fast data transfer rate between and within chips. Now, we can achieve high speeds with the use of optics, and if we integrate the optical components on a single silicon chip then it would allow the data to flow at a higher speed. This paper reviews the growth of silicon photonics compared to last few years and a promising future of photonics. Intel Corporation first announced its major breakthrough in 2004 based on silicon photonics (Moore's law). They developed a modulator based on silicon that operates at 50GHz. It was 50 times faster than the previous invention. In 2008 Intel declared that the data transfer rate can be increased further up to 1Tbps (bits per second).compatibility of silicon photonics with CMOS fabrication offers great advantages, such as low cost, high volume integration.*

Keywords – *Silicon, Modulators, Mach-Zehnder, SiGe, Detectors, Lasers, Avalanche Photodetector*

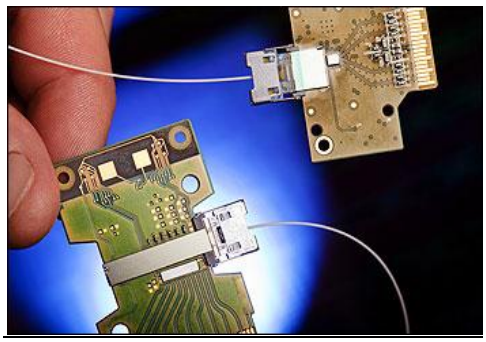
I. INTRODUCTION

Silicon photonics is the new technology of producing optical devices and circuits using silicon as the core material for the integration of optical and electronic components on single chip with standard CMOS (complementary metal oxide semiconductor) fabrication process. Individual Si or SiGe/Si components perform well at mid wave, long wave and very long –wave infrared (THz) ranges. The intrinsic band gap of silicon (1.1 eV) and its transparency at wavelengths (1270 nm to 1625 nm) typically favors for optical communication. It allows the integration of optical devices on silicon-on-insulator (SOI) wafers [1] which provides the route and manipulates the light. The present computers use copper wires to connect different components. These copper wires degrade the signal strength and have a limited maximum length. Intel's research achievement has replaced these connections with light weight optical fibres which can transfer more data over longer distances at higher speeds. Higher data rate can also be achieved through copper wires upto 10Gbps but with the increasing distance, speed starts decreasing.



Fig(1): Basic block diagram of optical link [2]

Above fig (1) shows the basic diagram of optical link. Firstly, it requires an on-chip laser source which provides the pumping mechanism. Recently researchers focus on Hybrid Silicon Laser, in which the silicon is bonded to a different semiconductor (InP/GaAs) which acts as the lasing medium. To carry out the data in form of optical signals we require a Modulator. Modulation is carried out through free carrier plasma dispersion effect in photonics. Silicon waveguide provides the optical medium between different chip and same chip. At the receiver end metal semiconductor junctions based photodetectors integrated into silicon waveguides for detection of light and converting back the data into electronic signals. In order to get the higher data rate over long distances, one possible solution is to use optical interconnections with silicon as an optical medium. In electrical medium signal information carried by the electrons but in optical medium information carried by the photons. Comparison to the electrons, photons has no charge and zero rest mass therefore photons can travel at higher rate i.e. speed of light and better stability. Therefore it is desirable to replace electrical circuits with optical circuits



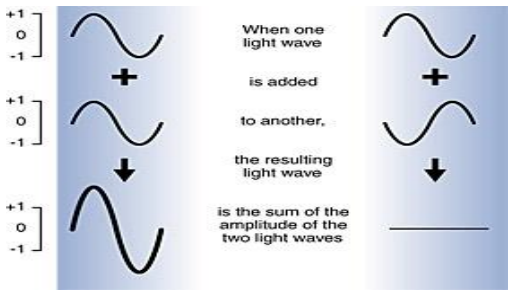
Fig(2): 50 Gbps silicon photonics transmit module [3]

Above fig(2) explains the 50Gbps optical link module generated by the Intel group (2007) [3]. Silicon chip presence at the centre of the board (lower side) provides laser light which travels through optical fibre, where a second silicon chip (upper side) detects the data on the laser and converts it back into an electrical signal. Optoelectronic systems integration on silicon on substrate (SOI) realizable with complementary metal-oxide semiconductor (CMOS) technology, allows low cost and large scale manufacturing for Si photonic devices. Further work going on the development of integration methods of low loss waveguides, high quality resonators, high speed modulators and optically pumped lasers on silicon chip for low cost and power consumption effective photonic circuits. This highly combination of Photonic technology fulfils the demands of numerous fields such as communication, computing, imaging and sensing. The most recent advancement involves encoding of high-speed data on an optical beam. This paper describes the photonic structures that are liable for Si based optic-electro communication link.

II. HIGH-SPEED SILICON MODULATION

An Optical Modulator is a device which is used to modulate a beam of light with respect to an information signal. Various factors decide the performance of the modulators: (1) modulation depth (2) modulation speed (3) bandwidth. In ideal cases, we prefer high modulation speed, large bandwidth and low power consumption.

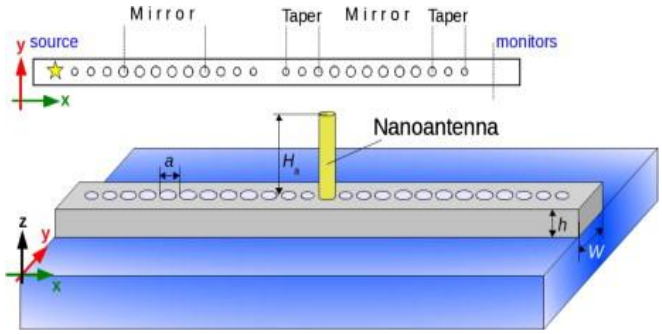
.



Fig(3): Modulation process of wave [4]

Fig (3) illustrates the modulation process where two wavelengths are combined. If two sine waves are perfectly matched (sync) then they are added together, resulting sine wave has twice the amplitude of the individual waves. In contrast, when two waves are completely out of sync, then resulting wave has no amplitude because two light waves cancel each other out [4]. Silicon is not considered to exhibit good electro-optic effect that could enable the modulation process but it favors the fabrication of various micro-optical devices. Silicon has transparency to infrared communication wavelength and has high refractive index which allows the miniaturization of photonic devices. In silicon integrated circuits modulation is come through by free carrier plasma dispersion effect.

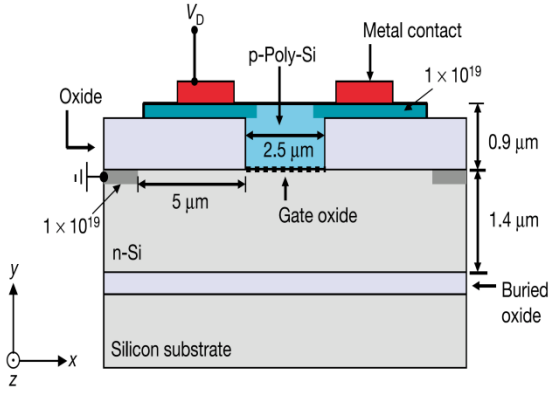
Various constraints such as bandwidth, narrow waveguide, modulation efficiency, low power consumption is kept in mind while integrating optical modulators on silicon substrate. Keeping in that mind recently, researchers reduced the waveguide core size from the micrometer range, producing silicon nanowires with a height and width of about 500nm and less [5].



Fig(4): Illustrates the silicon-on-insulator nano-beam cavity with a plasmonic nano-antenna.

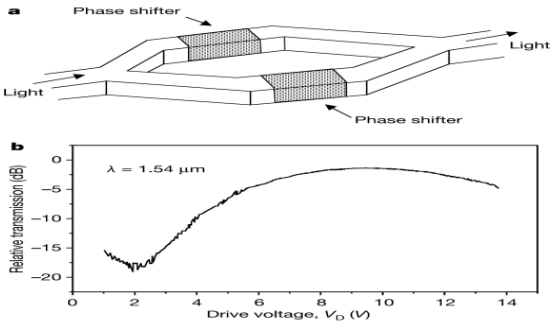
The cavity is formed by two tapered mirrors consisting of air holes in silicon rectangular nanowire (gray region, width W = 500 nm and height h = 260 nm) with a silica substrate (blue region). The distance between the holes of the mirrors is a = 375 nm and the radius of the holes is r = 0.28a. The radii of the holes in the tapers are 0.24a, 0.22a and 0.19a, respectively. The diameter and the height of the gold nano-antenna are D = 135 nm and Ha=500 nm, respectively [6]. Previously, we can achieve the speed up to ~10Gbps through silicon modulators but with the increasing demand of fast data rate we have to use the efficient and fast modulators. For high speed (>10Gbps), optical modulators based on electro-optic materials such as III-V semiconductors or LiNbO3 [7] are used in recent years. With these modulators, one can achieve modulation upto 50Gbps. It is not easy to achieve the high modulation because of crystalline structure of silicon, as it does not exhibit the linear electro-optic Pockels[8] effect and has a very weak Franz-Keldysh effect, Kerr effect [9,10] at optical communication wavelengths of 1.3 μm and 1.6 μm. Silicon also has high thermal coefficient (~2E-4)[11] which is not suitable for high speed operation of modulators. Recently, it has been shown that strained silicon (Ge/SiGe) structures have strong electro-optic effect due to quantum-confined Stark effect[12] making it desirable for optical modulation. Combining, modulators with recently developed hybrid silicon lasers, one could create a single chip that can transmit data at 1Tbps data rate. For fast optical modulation, free carrier plasma dispersion effect is considered, which is related with the concentration of free carriers in a semi-conductor, which changes both real and imaginary part of the refractive index. Change in the refractive index due to plasma dispersion effect can be used for modulation in two ways (1) Phase shifter /modulator (2) Resonant structures. Phase shifter /phase modulation configuration depends upon MachZehnder interferometer (MZI) which is used to determine the relative phase shift from the variations of refractive index in one or both arm of wave-guide. MZI configuration describes the two optical modulators structures based on carrier depletion and carrier accumulation mode.

III. Metal-oxide semiconductor (MOS) Capacitor-Based Silicon Modulator



Fig(5): Cross-section view of the MOS capacitor-based on silicon optical modulator.

In case of this modulator, modulation is achieved by driving the MOS capacitor in the accumulation mode of operation. Positive voltage $V_{D\ applied}$ on the gate, which forms the depletion and inversion layer across the waveguide region. Due to this positive voltage large amount of silicon charge carriers accumulate in the depletion region which changes the refractive index. Change in refractive index further changes the phase shift and attenuation in one or both arms of interferometer.

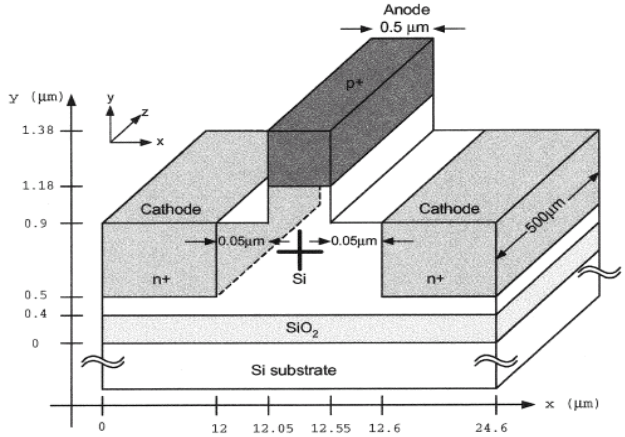


Fig(6): a) Phase shifter in both arms of MZI, b) Drive voltage change with relative transmission

When light enters into MZI it will split up into two beams in both arms of the interferometer. Phase shifter changes the relative phase shift between the two beams and modulation is achieved on the output of interferometer**.** Changes in the phase shift and attenuation in the arms of interferometer generated from the plasma dispersion effect. MOS capacitor also improves the modulation efficiency. Speed of charge carriers in MOS capacitor is independent upon minority carrier's lifetime but depends upon the device resistance and capacitance. The optical loss of the phase shifter in high-speed section is 10 dB/cm and 5.2 dB/cm for the low-speed sections. MZI modulator gives a total insertion loss of 19 dB consisting of 9 dB coupling loss and 10 dB on-chip loss. Out of 10 dB on-chip loss, 3.5 dB is due to the high-speed (RF) sections, 2.5 dB due to the low-speed (bias) sections[13]. Phase efficiency can be further be improved by reducing the waveguide dimensions and reducing the gate dielectric width. First breakthrough in MOS capacitor shown by Intel in 2004, with modulation bandwidth exceeding 1GHz[14].

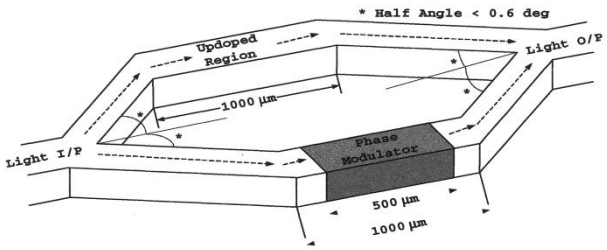
IV. P-N Or P-I-N BASED SILICON MODULATOR

Modulation is achieved in these modulators by operating the p-n diode[15] in reverse bias and in the carrier depletion mode. It is also based on MZI configuration and improvement of the MOS capacitor based modulators. Because of the reverse bias it requires low input power and independent of minority carrier life time i.e. no speed problem. But its main drawback is its low modulation depth which occurs due to overlap between the active electrical region and optical region.



Fig(7): p-i-n Si Optical Modulator

In this modulator, the dielectric layer between two doped regions replaced with the p-n junction. This structure arrangement reduces the modulator capacitance hence improved the bandwidth. The concentrations of n type are higher than the p type results carrier depletion mainly occurs in p type. It leads to better phase modulation efficiency because the hole density change results in a larger refractive index change as compared to the electron density change as indicated by[16]. Phase shift provided by the reverse biased p-n junction embedded in MZI arms shown in fig(8)

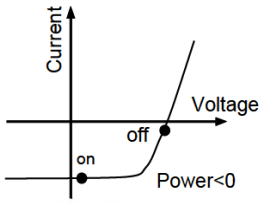


Fig(8): Phase modulator embedded in MZI arms.

The first carrier depletion-based silicon modulator built by Gardes et al. in 2005[17], bandwidth of 50GHz. After few years, Basak et al. in Intel came through with modulator (p-n diode based) with data rate of 40 Gbps and bandwidth of 30GHz[18,19]. Although carrier depletion modulators have higher speed operation but it requires improvement in the modulation depth.

V. Energy Harvesting

Sasan Fathpour and Bahram Jalali gave us another idea regarding optical modulators i.e. energy harvesting in silicon modulators. It is based on the consumption of pump energy that is being lost in two photon absorption (TPA) into some useful electrical power. For TPA to happen strong optical intensity light source is required. Modulation is carried out through electrical losses caused by TPA[20].The input optical power and wire cross-section are kept in such that optical intensity is high enough to induce the TPA. TPA induced free carriers give us V-I characteristics similar to the solar cell. These devices give us electrical power whenever biased in the fourth quadrant.



Fig(9): V-I characteristics in the fourth quadrant with negative power dissipation [20]

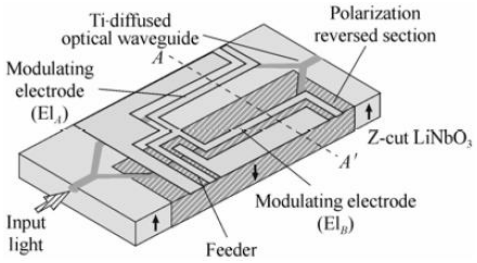
Above fig(9) shows that whenever device in fourth quadrant electrical power generated and device performs as a modulator otherwise energy is dissipates in the form of heat. Energy harvesting is significant for VLSI industry where high number of devices integrate on single chip resulting large dissipation of heat.

VI. Electro-absorption modulators

Electro absorption modulators are a semiconductor device that modulates the intensity of laser beam via electric voltage. Principle of operation based on the Franz–Keldysh effect. Compared to plasma dispersion-based modulators, absorption type modulators have high operating speed, low power dissipation and low operating voltage. Through these modulators a modulation bandwidth of 10 GHz can be achieved. Recently, it has been shown that strained silicon (GeSi)[21] or Ge/GeSi[22] multiple quantum wells also exhibits linear electro-optic refractive index modulation.

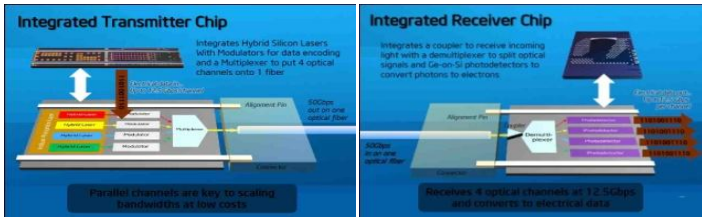
VII. Hybrid Silicon Modulator

Efficient modulators should have large bandwidth, high speed operation and good modulation efficiency. Modulators based on MZI have bandwidth close to 100nm while electro-absorption modulators are less than 30nm. Therefore for high speed operation, hybrid silicon modulators are developed, where the III-V material epitaxial layer grow on silicon-on-insulator (SOI) wafer by wafer bonding techniques[23]. LiNbO3 is the most widely used material as electro-optic modulators. It has high electro-optic coefficient and mature fabrication technology making it suitable for modulation applications. Previously, LiNbO3 modulators have modulation speed of 20Gbps but now ultrafast LiNbO3 of 60Gbps have been developed by titanium (Ti) diffusing [7].



Fig(10): LiNbO3 modulator with Ti diffusing[24]

Input signal is provided by the feeder which is further divided into two signals on the modulating electrodes E1A and E1B with same phases and amplitude. Phase shift provided by the arms of MZ interferometer in opposite direction. These Modulators were fabricated by the polarization reversal technique[24,25,26]. Modulators based on MachZehnder (MZ) or Electro-Absorption (EA) provides data rate higher than 10Gbps whereas MachZehnder structures on Lithium-Niobate (LiNbO3) give us data rate up to 60Gbps.To achieve higher data rate various new modulations format such as OPSK (Quadrature Phase Shift keying), DQPSK (Differential Quaternary Phase-Shift Keying)[27] are under progress. Recently, these formats based LiNbO3 modulator give us 80Gbps data rate. Currently, researcher's works on the multi-level modulation format based on MQAM (Quadrature-Amplitude Modulation) applied on single carrier to achieve the data rate of 100Gbps or beyond. Another approach supporting the higher data rate is the division of transmitted signal into sub carriers. These sub carriers are aligned orthogonally with the same symbol rate. Recently, Intel has also demonstrated 50Gbps optical link and their future work of achieving 1Tbps based on the sub carrier's modulation[3].



(a) (b)

Fig(11): Showing (a) Transmitter (b) Receiver with wavelength division multiplexing

Transmitter chip includes four hybrid silicon lasers (InP), generates four different colors of continuous-wave laser light or wavelengths. Wavelength division multiplexing (WDM) increases the number of channels in the fiber which increases the bandwidth. From the four hybrid silicon lasers, light beams travel into an optical modulator that encodes data on to them at 12.5Gbps. These four beams further combined to a single optical fibre for a total data rate of 50Gbps. At the receiver link, chip separates the four optical beams and directs them into photo detectors, which convert them back into electrical signals. Based on this model, research has been going on to achieve data rate of 100Gbps or further.

VIII. Photodetectors

Photodetectors is the semiconductor devices which convert the optical signals into electrical signals. Photodetectors should have high sensitivity in operating wavelength, high response speed, low noise, high reliability and low biasing across the detector. Photoconductor is the simplest and basic detector consists of single semiconductor slab across which electric field is applied. But it has high dark current noise and high response speed. Instead of this reverse biased pn junction has low dark current noise. Therefore photo detectors consist of p-n junction, works under a reverse bias voltage. Reverse bias voltage across the pn junction increases the reverse current and width of the depletion region. Thickness of the depletion region further increases the quantum efficiency but the response time increases. Photo detector performance mainly depends upon three factors (1) Responsivity (2) Dark current (3) Bandwidth.

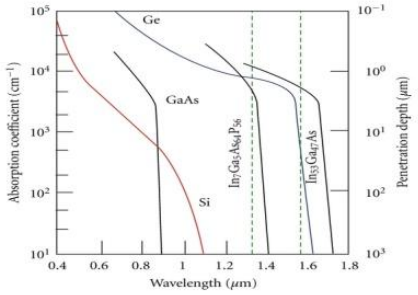
Dark current is the small electric current shown by the detector under no input light signals. It is also referred as reverse leakage current in detectors. Responsitivity determines the receiving efficiency of the detectors and calculated as

$$R = \eta \frac{e}{h\nu}$$

Where hν is the photon energy, η is the quantum efficiency, and e is the elementary charge. Efficiency can be increased by adding an intrinsic area in pn junction. When device operates in reverse biased then width of depletion region becomes equal to the intrinsic region. Therefore it increases the light absorption area and increases the quantum efficiency. Quantum efficiency can also be increased by using Anti-reflection coating on the surface. Bandwidth of the photo detectors depends upon junction capacitance and the generation of carriers across the junction. Bandwidth related to junction capacitance (Cj) given as

$$f_{BW} = 1 / (2 * \pi * R_{LOAD} * C_j)$$

With the increase in depletion width due to reverse, capacitance of the depletion region increases therefore bandwidth (response speed) increases. In most cases, detectors can have larger responsivity, but it will also increase both the dark current and the capacitance and hence reduces the sensitivity and the bandwidth, respectively[28]. Semiconductor materials, such as Si, Ge, and III-V material, like GaAs, InP, have been used for the manufacturing of hetrojunction photodetectors. Figure(12) compares the absorption coefficients of different semiconductors including Silicon. From this fig(12) it can be seen that Si and GaAs have very low absorption coefficient while Ge and some other III-V materials have considerable absorption coefficient



Fig(12): Absorption coefficient as a function of their wavelength. The green dashes mark typical wavelengths for telecommunications windows of 1310 and 1550 nm [29]

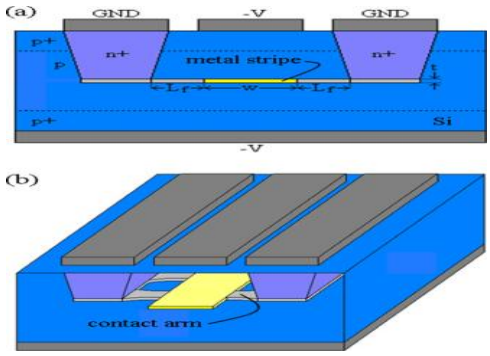
Silicon Detectors

From the fig(12) it is clear that Si exhibit transparency in the 1300–1550 nm operating wavelength range because of its low absorption coefficient it is not a suitable material for fibre-optic communication application. But it performs well under 1000nm, where band-to-band absorptions happen. To solve this problem, we can approach to three structures; Schottky structure, III-V heterogeneous detectors, or Ge detectors built on silicon substrates.

IX. Schottky Detectors

Schottky detectors are also known as metal semiconductor diode, consists of a thin metal layer (~ 100 Å) on a lightly doped p-type Si, forming a Schottky barrier at the material interface. Schottky barrier results the depletion layer between the metal and semiconductor junction also known as the absorption region. The barrier is usually 0.2–0.6 eV for p-Si, while the value is a bit smaller for n-Si with the same contact metal. Main advantages of Schottky detectors that it is majority carrier device, not dependent on minority carrier's lifetime therefore no delays problems. It has also high response speed (150GHz) and compatible with standard CMOS technology. But its drawback is the low quantum efficiency. Quantum efficiency can be improved by merging the dielectric waveguide using transparent conducting electrodes and employing surface Plasmon polaritons (SPP) [30,31]. Fig(11) shows below symmetric and asymmetric surface Plasmon Schottky contact detector. Scales et al. demonstrated symmetric structure with responsivity of 0.1 A/W and dark current of 21nA[32].

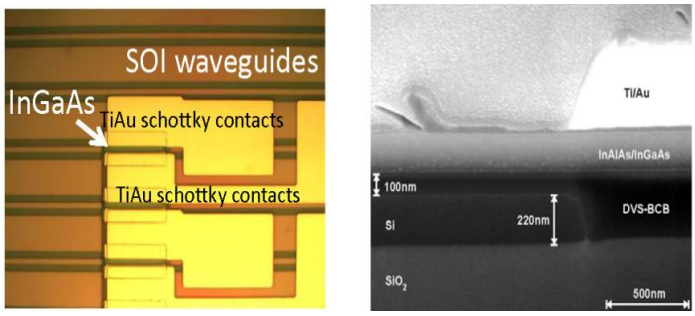
In the symmetric structure, fabrication process is complicated as it involves the realization of thin metal films buried in semiconductor in order to simply the fabrication process Berini`s group proposed the asymmetric Surface Plasmon polaritions photodetectors[33,34].



Fig(13): Scheme diagram of examples of (a) symmetric and (b) asymmetric surface Plasmon Schottky contact detector.

X. III-V Heterogeneous Detector

III-V semiconductors have direct band gap and compared to silicon III-V-based detectors have wider absorption bandwidth and relative low dark currents. But the lattice structures of III-V materials are different from the Si therefore these semiconductor materials cannot be directly grown on silicon. One solution to this problem is the heterogeneous integration of III-V layers on silicon substrate using advanced wafer bonding technique. In earlier times III-V detectors (GaAs,) are not compatible with CMOS technology approach. But in recent times some of them, like InGaAs/InP built on the SOI substrate using heterogeneous integration approach. This integration process achieved by the DVS-BCB adhesive die-to-wafer bonding process [35].In fig (12) metal-semiconductor-metal photo detector (MSM) using an InGaAs absorption layer were realized. Coupling is formed between the SOI waveguide and photo detector by the vertical directional coupling between the SOI waveguide and the lossy III-V waveguide mode. The dark current of the device was 3.0nA at a bias voltage of 5V, while the responsivity of detector is 1.0A/W at 1.55μm and a bias voltage of 5V[36].



(a) (b)

Fig(14): (a) Top view of the fabricated III-V metal-semiconductor-metal photodetectors integrated on an SOI waveguide platform (b) cross-section of the device.

XI. Germanium Detector built on Silicon

Ge has indirect bandgap structure similar to Si but its band gap energy is less (0.7ev versus 1.1ev for Si) compared to Si. Therefore it has higher absorption coefficient up to wavelength of 1550nm [29]. Today, most of the researchers work on Ge-on-Si photodetectors because of its high responsivity behavior in near IR wavelength, its optoelectronic properties and compatibility with CMOS fabrication. Ge/GeSi can be deposited by chemical vapour deposition process (CVD). In these photodetectors main difficulty comes in growing o epitaxial films of Ge on Si. Due to the high lattice constant of Ge 4% greater than the silicon, there is always mismatched in their structures. This will cause defects while growing Ge layers on Si such as dislocations in their densities arrangement and high power loss. These defects greatly affect the performance of the photodetectors. To overcome with this problem we have two different approaches (1) selectively growing of Ge on Si (2) two step method of growing Ge. In the selective approach we generally use Si3N4 as the dielectric mask through which Ge layers grow on the silicon. Deposition of layers of Ge is done through chemical vapour deposition (CVD) through the exposed areas develop over the dielectric mask[1,9]. In two step method Ge layers grow epitaxially on Si under controlled temperature. Initially, Ge layer of 30 nm grown on Si at a temperature of 320–360°C. In the second step continuous growth of Ge desired over the Si at a higher temperature greater than 600°C. Later on Loh et al. [37] demonstrated modified two step approach, which includes growing of ultrathin (2~30 nm) Low Temperature SiGe buffer layer prior to the deposition of Low Temperature Ge layer and High Temperature Ge layer.

There are two structures of Ge-on-Si p-i-n photo detectors (1) normal-incidence type (2) waveguide type. Waveguide based photo detectors is better than normal-incidence photo detectors on the basis of quantum efficiency and bandwidth. Waveguide detectors have smaller area than the incidence type, about ten times. There has been great progress going on Ge on Si photo detectors. Vivien et al. demonstrated the Ge detector with bandwidth of 42 GHz and responsivity at 1 A/W operating at 1550 nm wavelength[38]. Feng et al. also showed the zero bias Ge detector with 17.5GHz bandwidth[39]. CMOS integration process lowers down the bias voltage across the detectors, reduce the dark current, power absorption and operating voltage of chip. Recently Luxtera[40] and MIT[41] demonstrated Ge detectors of 130 nm and 180 nm respectively integrated with CMOS technology Currently, a zero bias Ge detector with 17.5 GHz has been demonstrated by. In 2010, Kang group reported Ge detector with high sensitivity at 4 dB and speed at 10Gb/s.

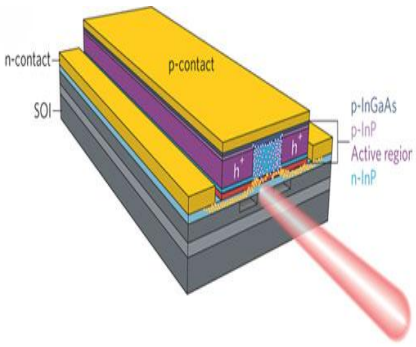
XII. Avalanche Photodetector

Avalanche photodiode detectors (APD) have better sensitivity than the p-i-n type detectors and exhibit larger gain bandwidth product. Ge-on-Si APD gives us better performances than the III-V detectors. In a avalanche based photodetectors number of charge carriers increases by 10-100 factor due to multiplication of charge carriers. Therefore it increases the sensitivity, bandwidth and helpul in long distance transmission. Progress work going on the development of low cost Ge-on-Si APD with data rate of 40Gbps. Recently the Intel-UCSB team has demonstrated a 340 GHz gain–bandwidth Ge–Si avalanche photodiode using CVD growth of Ge and Si layers at 850°C[42].

XIII. LIGHT EMITTERS AND LASER

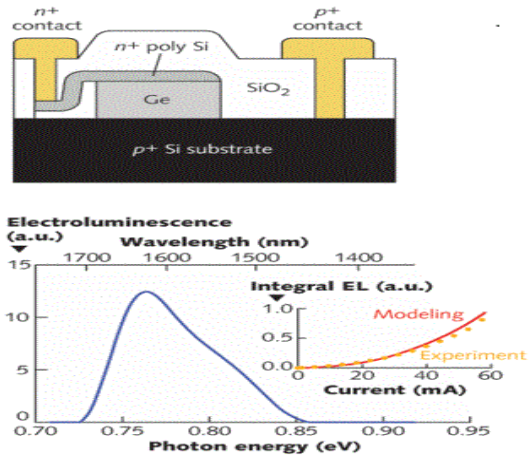
Si has an indirect bandgap structure which implies that radiative recombination requires emission or absorption of phonon. The involvement of phonon makes it very inefficient light emitter. Several progresses have been made in silicon based light emitters and laser sources. Initially researchers focus on the quantum dot lasers and fusing of impurities into silicon such as neodymium but these attempts are not useful. Later on researchers work on some other alternatives. In 2002[43] ,idea was proposed of using Stimulated Raman Scattering (SRS) in silicon waveguides.Two years later, Jalali's group demonstrated the first Si laser based on Raman Effect at UCLA. But it has free carrier absorption (FCA) and mean carrier lifetime problems which make it impossible to achieve the gain in silicon waveguides. FCA occurs when a material absorbs a photon and a carrier is excited from a filled state to an unoccupied state (in the same band). FCA is removed by integrate the waveguide structure into the intrinsic region of a PIN diode, which is reverse biased so that the carriers are attracted away from the waveguide. In 2005[44], Intel Corp. reported the first continuous-wave (CW) Si Raman laser based on reversed biased p-i-n Si waveguide structure. Si properties like high optical damage threshold, high thermal conductivity and large reflective index contributes the silicon as an excellent element for Raman crystal. Si Raman laser operate well in 1550nm near-IR band but it has one problem i.e. the loss induced by two-photon absorption (TPA) and Free-carrier absorption (FCA). TPA is a nonlinear loss mechanism, it occurs when two photons combine their energies to boost an electron in the valence band to the conduction band. Silicon Raman lasers are considered to be the ideal light sources because of its unmatched wavelength purity and the possibility of extending the lasing wavelength into the mid-infrared region. Researchers also found that the action of silicon nanocrystals on erbium ions give us a new base for silicon based optical amplifiers and laser sources. The main advantages of erbium-doped silicon is (1) easy to fabricate(2) CMOS compatible(3) wavelength suitable for optical communication(4) optical activity demonstrated. For improving the emitter efficiency multi layered structures of Er doped Si nanocrystals preferred. Efficiency is also dependent upon the doping concentration and device geometry. Main problem with erbium doped silicon is that it forms the Auger process because erbium is electrically active and its high pumping mechanisms.

There is another light source in which III-V-based materials (GaAs and InP) are grown epitaxially on silicon substrate. Compared to Si, GaAs and InP have direct bandgap structure which favors the laser mechanism. But the major problem is large lattice mismatch between III-V-based materials (GaAS and InP) and Silicon which results dislocation of densities of 10E8–10E10 cm–2 [45] respectively. Another disadvantage is III-V materials epitaxial growth methods are not compatible with CMOS processing results large power loss. Various approaches have been made to remove the dislocations such as direct wafer bonding, strained super lattices [46], divinyltetramethyl-disiloxanebenzocyclobutene (DVS-BCB) adhesive bonding[7].



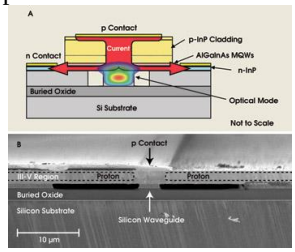
Fig(15): Injected carriers in iii–v epitaxial layers on top of a Si waveguide.

To overcome with the problem of III-V-based materials growth on silicon substrate, researchers now work on the optimized growth of III-V materials under low temperatures. One of the examples is InGaAs nanopillar (NP) laser compatible with CMOS processes. Another useful approach is Ge-on-Si (or SiGe-on-Si) epitaxial growth. Ge and Si both has mismatch in their lattice constant and thermal coefficients. Germanium has an direct band but it has the chances of radiative recombination because of its lower energy gap (0.8eV). High thermal coefficients of Ge also favor the radiative recombination because of reduction in tensile strain of is0.2–0.25% [47]. These are the basis of direct bandgap electroluminescence and optically pumped operation of Ge-on-Si lasers. MIT demonstrated the first Ge on Si continuous wave laser working at room-temperature, fabricated by epitaxial growth of 1.6μm * 0.5 μm Ge waveguides on Si[48].



Fig(16): Ge epitaxial growth on Si and the electroluminescence behavior.

Instead of growing single material on substrate researchers developed the new formation where combinations of several materials grown on silicon substrate simultaneously. Hybrid silicon laser is a semiconductor device where III-V group materials fabricated on silicon substrate. III-V group materials have direct band gap structure compared to Si therefore in hybrid approach we use the light emitting properties of III-V materials with mature fabrication technique. In 2006, first electrical driven Hybrid Silicon Laser was developed with collaboration between the University of California, Santa Barbara, and Intel Corporation[49]. These lasers do not require any optical input and it can directly fabricate on to silicon chip. In the hybrid laser unpatterned III-V wafer is bonded to a patterned silicon wafer with optical waveguides. Silicon waveguides are patterned before laser fabrication so no alignment is needed between the unpatterned III-V wafer and the patterned silicon wafer.

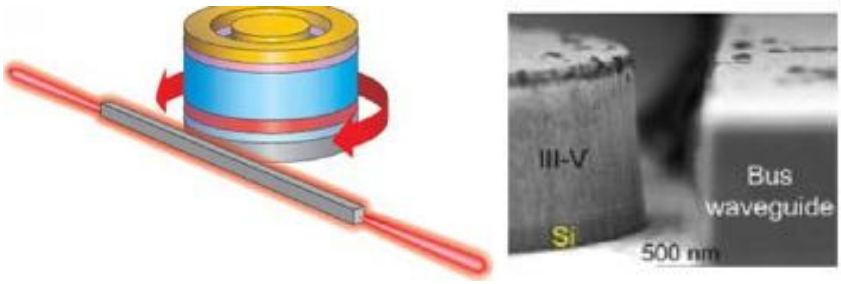


.

Fig(17): Hybrid Silicon Laser (a) and the scan view (b)

When light or electrical signals drives the hybrid laser then light emits from active layer of AlGaInAs quantum

wells in the III-V material couples into the silicon waveguide. In Hybrid Silicon Laser, III-V material used for optical gain/modulation and silicon used for optical wave-guiding medium. Gain can be varied by adjusting both the height and the width of the silicon waveguide and it is dependent upon the overlapping between the hybrid-waveguide mode and the quantum-well region. Using this approach, a variety of discrete active devices such as distributed feedback (DFB), distributed Bragg reflector (DBR), ring fabry-perot and AWG lasers have been fabricated. Currently, Hybrid III-V on Si technology is the most advanced photonic integrated circuits on Si. Researchers now working on the new structures of hybrid laser e.g. Ring or disk geometries for hybrid laser. First electrically pumped hybrid Si micro-ring laser was demonstrated by Liang et al.[50].



(a) (b)

Fig(18): (a) Hybrid micro-ring laser with a Si bus waveguide. (b)SEM image of resonator sidewall and bus waveguide

XIV. APPLICATIONS AND FUTURE ASPECTS

Silicon photonics, extremely powerful technology in the recent trades, fulfills the demands for higher speed, efficiency, and low power consumption at lower costs. In silicon technology, various opto-electronics devices integrate on a single substrate connected with each other using narrow waveguides. Such circuits could be used to establish high speed transmission, increased bandwidth, reduced power consumption and decreasing latency problems. It can potentially increase the bandwidth capacity by providing micro-scale, ultra low power devices. Optical data transmission increases the data rate and eliminates the problem of electromagnetic interference. With these data rates one could imagine the videoconferencing with a high resolution that the actors or family members appear to be in the room with you. Optical link can transfer data over longer distances and faster than today's copper technology; up to 50 GB of data per second. Recently, Intel Corporation launched their optical link connection operating 100Gbps[51]. Fujitsu Laboratories recently developed four wavelength integrated silicon laser for optical transrecievers. Work is in the progress for the new optic-interconnection of external devices to PCs. Intel introduced Light Peak, 2009 which replaces the USB and communicates the data at up to 10Gbps. Advancement in the field of silicon nanophotonic technologies leads new ideas in future computing systems and their architectures.

XV. DRAWBACKS

(1) Silicon's main drawback is its poor optical emission efficiency due to its indirect band-gap. Silicon has inefficient electroluminescence; therefore it suffers from light sources. For these reasons, R&D focuses on the development of practical silicon light sources. Lasers are preferred as a light source but it requires high pumping mechanism. Significant progress has been made in the field of hybrid silicon laser which allow low cost and provides efficient laser light. (2) Optical connections require precise alignment therefore it requires improved alignment technologies.(3) Heat dissipated by the laser source on chip.(4) Band-gap of silicon is larger than desirable making it impossible to detect light in the telecom spectral range 1.5 and1.3μm.(5) Current fabrication processes are localized for making pure electronics devices. Integrating photonic components on the same chip requires additional fabrication steps that may interrupt the standard process and affect the performance of the electronic components. (6) Integration of components increasing day by day on chip therefore power dissipation of components also increases across the chip.

REFERENCES

[1] R. A. Soref.: 'The past, present, and future of silicon photonics', IEEE J. Sel. Top. Quantum Electron, 12, (6), pp. 1678– 1687, 2006.
[2] Ali Okey, Thesis, SI-GE Photodetection Technologies for integrated optoelectronics, partialfulfilment of doctraldiss, Stanford University, 2007.
[3] Http://www.intel.com/pressroom/archive/releases/2010/20100727comp_sm.htm
[4] Dr. Mario Paniccia, Victor Krutul, Sean Koehl, Intel Unveils Silicon Photonics Breakthrough: High-Speed Silicon Modulation, Technology@Intel Magazine, 2004.
[5] Frederic Gardes, GoranMashanovich, GrahamReed,"Evolution of optical modulation in silicon-on-insulator devices", SPIE Newsroom. DOI: 10.1117/2.1200712.0985, 2007.
[6] Ivan S. Maksymov,"Optical switching and logic gates with hybrid plasmonic–photonic crystal nanobeam cavities "Physics Letters A 375 ,2011 918–921.

[7] K. Noguchi, "Ultra-high-speed LiNbO3 modulators," Journal of Optical and Fiber Communications Reports, vol. 4,no. 1, pp. 1–13,2007.

[8] R.S Jacobsen,K.N Anderson ,P.I.Borel,ETAl.,"Stranded Silicon as a New Electrooptic material, "Nature, Vol.441, No.7090.PP.199-202.2006.

[9] R. A. Soref and J. P. Lorenzo, "All-silicon active and passive guided-wave components for $\lambda = 1.3$ and 1.6 μm," IEEE Journal of Quantum Electronics, vol. 22, no. 6, pp. 873–879, 1986.

[10] R. A. Soref and B. R. Bennett, "Electro optical effects in silicon," IEEE Journal of Quantum Electronics, vol. 23, No. 1, pp. 123–129,1987.

[11] G. T. Reed and A. P. Knights, Silicon Photonics: An Introduction, John Wiley & Sons, Chichester, UK, 2004.

[12] Y.-H. Kuo, Y. K. Lee, Y. Ge, et al., "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," Nature,vol. 437, no. 7063, pp. 1334–1336, 2005.

[13] L. Liao, A. Liu, R. Jones, et al., "Phase modulation efficiency and transmission loss of silicon optical phase shifters," IEEE Journal of Quantum Electronics, vol. 41, no. 2, pp. 250–257, 2005.

[14] Ansheng Liu1, Richard Jones1, Ling Liao1, Dean Samara-Rubio1, Doron Rubin2, Oded Cohen2, Remus Nicolaescul & Mario Paniccia1,"A high-speed silicon optical modulator based on a metal–oxide–semiconductor capacitor", Nature427, 615-618,2004.

[15] C. K. Tang, G. T. Reed, "Highly efficient optical phase modulator in SOI waveguides," Electron. Lett. 31, 451-452,1995.

[16] Http://www.intel.com/products/processor/core2XE/qc_prod_brief.pdf.

[17] F. Y. Gardes, G. T. Reed, N. G. Emerson, and C. E. Png, "A sub-micron depletion-type photonic modulator in silicon On insulator, "Optics Express, vol. 13, no. 22, pp. 8845–8854, 2005.

[18] L. Liao, A. Liu, D. Rubin et al., "40 Gbps silicon optical modulator for high-speed applications," Electronics Letters, vol. 43, no. 22,pp. 1196–1197, 2007.

[19] A. Liu, L. Liao, D. Rubin et al., "Recent development in a high-speed silicon optical modulator based on reverse-Biased pn diode in a silicon waveguide," Semiconductor Science and Technology, vol. 23, no. 6, Article ID 064001, 2008.

[20] S. Fathpour and B. Jalali, "Energy harvesting in silicon optical modulators," Optics Express, vol. 14, no. 22, pp. 10795–10799, 2006.

[21] J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L. Kimerling, and J. Michel, "Waveguide-integrated, ultralow-energy GeSi electroabsorption modulators," Nat. Photonics 2(7), 433–437 , 2008.

[22] Y. H. Kuo, Y. K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. Miller, and J. S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," Nature 437(7063), 1334–1336 , 2005.

[23] D. Liang, G. Roelkens, R. Baets, and J. Bowers, "Hybrid integrated platforms for silicon photonics," Materials 3(3), 1782–1802, 2010.

[24] Ichikawa J, Oikawa S, Yamamoto F, et al. Zero chirp broadband Z-cut LiNbO3 optical modulator using polarization reversal and branch electrode. Proceedings of 29th Optical Fiber Communication Conference (OFC'04): Vol 1, Feb 23-27, 2004, Los Angeles, CA, USA.Piscataway, NJ, USA: IEEE, 2004.

[25] Courjal N, Porte H, Martinez A, et al. "LiNbO3. Mach-Zehnder modulator with chirp adjusted by ferroelectric domain inversion" IEEE Photonics Technology Letters, 14(11): 1509−1511, 2002.

[26] Murata H, Kaneda K, Yamamoto S." Low-power and low-chirp guided-wave electrooptic intensity modulator by use of domain-inverted structure" Proceedings of Conference on Lasers and Electro-Optics (CLEO'03), Baltimore, MD, USA. Piscataway, NJ, USA: IEEE, 2003: 1008−1010, 2003.

[27] Lu G W, Miyazaki T, Ichikawa J, et al. High-speed DQPSK transmitter using a monolithically integrated quad Mach-Zehnder IQ modulator driven at quarter bit-rate. Proceedings of the 34th European Conference on Optical Communication (ECOC'08), Sep 21−25, 2008, Brussels, Belgien. Piscataway, NJ, USA: IEEE, 2008.

[28] Zhou Fang, Ce Zhou Zhao," Recent Progress in Silicon Photonics: A Review", ISRN Optics Volume 2012 , Article ID 428690,2012.

[29] B. Jalali, M. Paniccia, and G. Reed, "Silicon photonics," IEEE Microwave Magazine, vol. 7, no. 3, pp. 58–68, 2006.

[30] Reather, H., 1988, Surface Plasmons on Smooth and Rough Surfaces and on Gratings. Springer, Berlin, 1988.

[31] Ebbesen, W., Lezec, H. J., Ghaemi, H. F., Thio, T., Wolff, P. A., Extraordinary optical transmission through sub-wavelength hole arrays, Nature, 391, 667-669,1997.

[32] C. Scales, I. Breukelaar, and P. Berini, "Surface-plasmonSchottky contact detector based on a symmetric metal stripe in silicon, "Optics Letters, vol. 35, no. 4, pp. 529–531, 2010.

[33] Akbari A., and Berini, P., Schottky contact sur-face-plasmon detector integrated with an asymmetric metal stripe waveguide, Appl. Phys. Lett., 95(2), 021104,2009.

[34] A. Liu, R. Jones, L. Liao, D. Samara-Rubio, D. Rubin, O. Cohen, R. Nicolaescu, and M. Paniccia, "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," Nature 427(6975), 615–618, 2004.

[36] J. Brouckaert et al., "Compact InAlAs/InGaAs metal semiconductor-metal photodetectors coupled with SOI waveguides, submitted for publication in Photonics Technology Letters".

[37] T. H. Loh, H. S. Nguyen, C. H. Tung et al., "Ultrathin low temperature SiGe buffer for the growth of high quality Ge epi layer on Si(100) by ultrahigh vacuum chemical vapor deposition," Applied Physics Letters, vol. 90, no. 9, Article ID 092108, 2007.

[38] L. Vivien, J. Osmond, J. M. Fédéli et al., "42 GHz p.i.n germanium photodetector integrated in silicon-on-insulator waveguide,"Optics Express, vol. 17, no. 8, pp. 6252–6257, 2009.

[39] D. Feng, S. Liao, P. Dong et al., "High-speed Ge photodetector monolithically integrated with large cross-section silicon-on-insulator waveguide," Applied Physics Letters, vol. 95, no. 26, Article ID 261105, 2009.

[40] G. Gunn, "CMOS Photonics for high-speed interconnects," Journal Microelectronics, vol. 26, pp. 58–66, 2006.

[41] M. Beals, J. Michel, J. F. Liu et al., "Process flow innovations for photonic device integration in CMOS," in Silicon Photonics III, vol.6898 of Proceedings of SPIE, San Jose, Calif, USA, January 2008.

[42] Kang, Y.M., Liu, H.D., and Morse, M., et al.: "Monolithic germanium/ silicon avalanche photodiodes with 340 GHz gain-bandwidth product", Nature Photonics, 3, (1), pp. 59–63, 2009.

[43] Jalali, B., Raghunathan, V., and Dimitropoulos, D., et al.: "Raman based silicon photonics", IEEE J. Sel. Top. Quantum Electron, 12, (3), pp. 412–421, 2006.

[44] Rong, H.S., Jones, R., and Liu, A.S., et al.: "A continuous-wave Raman silicon laser", Nature, 433, (7027) ,pp. 725–728, 2005.

[45] Kawanami, H. ,"Hetero-epitaxial technologies of III–V on Si". Sol. Energy. Mat. Sol. C. 66, 479–486, 2001.

[46] Masafumi, Y., Mitsuru, S. & Yoshio, I. "Misfit stress dependence of dislocation density reduction in GaAs films on Si substrates grown by strained-layer super lattices". Appl. Phys. Lett. 54, 2568–2570, 1989.

[47] Liu, J. et al. "Tensile-strained, n-type Ge as a gain medium for monolithic laser integration on Si" Opt. Express 15, 11272–11277, 2007.
[48] Liu, J., Sun, X., Camacho-Aguilera, R., Kimerling, L. C. & Michel, J. Ge-on-Si laser operating at room temperature. Opt. Lett. 35, 679–681, 2010.
[49] Fang, A.W., Park, H., and Cohen, O., et al.: "Electrically pumped hybrid AlGaInAs-silicon evanescent laser", Opt. Express, 14, (20),pp. 9203–9210, 2006.
[50] D. Liang, M. Fiorentino, T. Okumura, H.H. Chang, D.T. Spencer, Y.H. Kuoetal."Electrically-pumped compact Hybrid silicon microring lasers for optical interconnects" Optics Express, pp. 20355–20364, 2009.
[51] D. McGrath, "Intel is developing optical chip-to-chip interconnects," Electronic Engineering Times, vol. 1556, p. 39, 2009.

Analysis in Theory and Applications
Anal. Theory Appl., Vol. **29**, No. 1 (2013), pp. 86-96

DOI: 10.4208/ata.2013.v29.n1.9

Simultaneous Approximation for Szász-Mirakian-Stancu-Durrmeyer Operators

Vijay Gupta[1,*], Naokant Deo[2] and Xiaoming Zeng[3]

[1] *School of Applied Sciences, Netaji Subhas Institute of Technology, Sector 3 Dwarka, New Delhi, 110078, India*
[2] *Department of Applied Mathematics, Delhi Technological University, Bawana Road, Delhi, 110042, India*
[3] *Department of Mathematics, Xiamen University, Xiamen 361005, Fujian, China*

Received 28 March 2012

Abstract. The aim of this work is to generalize Szász-Mirakian operator in the sense of Stancu-Durrmeyer operators. We obtain approximation properties of these operators. Here we study asymptotic as well as rate of convergence results in simultaneous approximation for these modified operators.

Key Words: Szász-Mirakian-Stancu-Durrmeyer operator, simultaneous approximation, asymptotic, rate of convergence.

AMS Subject Classifications: 41A20, 41A25, 41A35

1 Introduction

Let α and β be two non-negative parameters satisfying the condition $0 \leq \alpha \leq \beta$. For any nonnegative integer n,

$$f \in C[0,\infty) \to S_n^{(\alpha,\beta)} f,$$

the Stancu type Szász-Mirakian-Durrmeyer operators are defined by

$$S_{n,r}^{(\alpha,\beta)}(f,x) = n \sum_{k=0}^{\infty} s_{n,k}(x) \int_0^{\infty} s_{n,k+r}(t) f\Big(\frac{nt+\alpha}{n+\beta}\Big) dt, \tag{1.1}$$

where

$$s_{n,k}(x) = e^{-nx} \frac{(nx)^k}{k!}.$$

*Corresponding author. *Email addresses:* vijaygupta2001@hotmail.com (V. Gupta), dr_naokant_deo@yahoo.com (N. Deo), xmzeng@xmu.edu.cn (X. M. Zeng)

For $\alpha=\beta=0$ these operators become the well known Szász-Mirakian-Durrmeyer operators

$$S_n^{(0,0)}(f,x)=S_n(f,x)$$

introduced by Mazhar and Totik [3]. In [1] the author established some direct results in simultaneous approximation for this special case. Gupta et al. [2] estimated the rate of convergence for functions having derivatives of bounded variation for this special case $\alpha=\beta=r=0$. Also for this special case [4] estimated the rate of convergence for the Bézier variant of Szász-Mirakian-Durrmeyer operators.

The purpose of this paper is to study approximation properties of the Stancu type Szász-Mirakian-Durrmeyer operators. We give the rate of convergence and Voronovskaya type asymptotic result for the same operators.

2 Basic results

In this section we establish a recurrence formula for the moments.

For simultaneous approximation, we need the following form of the operators (1.1)

$$S_{n,r}^{(\alpha,\beta)}(f,x)=n\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}s_{n,k+r}(t)f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt.$$

Lemma 2.1. *For $n,m\in\mathbf{N}\cup\{0\}$, $0\le\alpha\le\beta$, let us consider*

$$\mu_{n,m,r}^{(\alpha,\beta)}(x)=S_{n,r}^{(\alpha,\beta)}\big((t-x)^m,x\big)=n\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}s_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt,$$

we get

$$\mu_{n,0,r}^{(\alpha,\beta)}(x)=1,\qquad \mu_{n,1,r}^{(\alpha,\beta)}(x)=\frac{\alpha+r+1-\beta x}{n+\beta},$$
$$\mu_{n,2,r}^{(\alpha,\beta)}(x)=\frac{\beta^2x^2+2(n-\alpha\beta-\beta-\beta r)x+(\alpha+r+1)(\alpha+r+2)-\alpha}{(n+\beta)^2},$$

and

$$(n+\beta)\mu_{n,m+1,r}^{(\alpha,\beta)}(x)=x[\mu_{n,m,r}^{(\alpha,\beta)}(x)]'+(m+\alpha+r+1-\beta x)\mu_{n,m,r}^{(\alpha,\beta)}(x)$$
$$+m\Big(\frac{2(n+\beta)x-\alpha}{n+\beta}\Big)\mu_{n,m-1,r}^{(\alpha,\beta)}(x). \tag{2.1}$$

Proof. By simple calculation we can easily obtain

$$xs'_{n,k}(x)=(k-nx)s_{n,k}(x).$$

We have from the definition of $\mu_{n,m}^{(\alpha,\beta)}(x)$

$$\begin{aligned}x[\mu_{n,m,r}^{(\alpha,\beta)}(x)]' =& n\sum_{k=0}^{\infty} xs'_{n,k}(x)\int_0^{\infty} s_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt\\ &-nmx\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} s_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^{m-1} dt\\ =& n\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty}(k-nx)s_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt - mx\mu_{n,m-1,r}^{(\alpha,\beta)}(x).\end{aligned}$$

Now

$$\begin{aligned}&x\Big[[\mu_{n,m,r}^{(\alpha,\beta)}(x)]'+m\mu_{n,m-1,r}^{(\alpha,\beta)}(x)\Big]\\ =& n\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} ts'_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt + n^2\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} ts_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt\\ &-n(nx+r)\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} s_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt.\end{aligned}$$

Putting

$$t=\Big(\frac{n+\beta}{n}\Big)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)-\frac{\alpha}{n}+\Big(\frac{n+\beta}{n}\Big)x,$$

we have

$$\begin{aligned}&x\Big[[\mu_{n,m,r}^{(\alpha,\beta)}(x)]'+m\mu_{n,m-1}^{(\alpha,\beta)}(x)\Big]\\ =& n\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} s'_{n,k+r}(t)\Big(\frac{n+\beta}{n}\Big)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^{m+1} dt\\ &-n\Big(\frac{\alpha-(n+\beta)x}{n}\Big)\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} s'_{n,k+r}(t)\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt\\ &+n^2\sum_{k=0}^{\infty} s_{n,k}(x)\int_0^{\infty} s_{n,k+r}(t)\Big(\frac{n+\beta}{n}\Big)\Big\{\frac{nt+\alpha}{n+\beta}-\frac{\alpha-(n+\beta)x}{n}-x\Big\}\Big(\frac{nt+\alpha}{n+\beta}-x\Big)^m dt\\ &-(nx+r)\mu^{(\alpha,\beta)}{}_{n,m,r}(x)\\ =& -\Big(\frac{n+\beta}{n}\Big)(m+1)\mu_{n,m,r}^{(\alpha,\beta)}(x)\Big(\frac{n}{n+\beta}\Big)-\Big\{\frac{\alpha-(n+\beta)x}{n}\Big\}(-m)\mu_{n,m-1,r}^{(\alpha,\beta)}(x)\Big(\frac{n}{n+\beta}\Big)\\ &+n\Big(\frac{n+\beta}{n}\Big)\mu_{n,m+1,r}^{(\alpha,\beta)}(x)-n\Big\{\frac{\alpha-(n+\beta)x}{n}\Big\}\mu_{n,m,r}^{(\alpha,\beta)}(x)-nx\mu_{n,m,r}^{(\alpha,\beta)}(x)\\ =& -(m+1)\mu_{n,m,r}^{(\alpha,\beta)}(x)+m\Big\{\frac{\alpha-(n+\beta)x}{n+\beta}\Big\}\mu_{n,m-1,r}^{(\alpha,\beta)}(x)+(n+\beta)\mu_{n,m+1,r}^{(\alpha,\beta)}(x)\\ &-\{\alpha-(n+\beta)x\}\mu_{n,m,r}^{(\alpha,\beta)}(x)-(nx+r)\mu_{n,m,r}^{(\alpha,\beta)}(x).\end{aligned}$$

Hence

$$(n+\beta)\mu_{n,m+1,r}^{(\alpha,\beta)}(x)=x[\mu_{n,m,r}^{(\alpha,\beta)}(x)]'+(m+\alpha+r+1-\beta x)\mu_{n,m,r}^{(\alpha,\beta)}(x)$$
$$+m\left\{\frac{2(n+\beta)x-\alpha}{n+\beta}\right\}\mu_{n,m-1,r}^{(\alpha,\beta)}(x).$$

This completes the proof. □

Remark 2.1. From Lemma 2.1, for $n\geq\beta^2+(\alpha+r)^2+3r+2\alpha+2$ and any $x\in(0,\infty)$, we have

$$\mu_{n,2,r}^{(\alpha,\beta)}(x)\leq\frac{(x+1)^2}{n+\beta}.$$

Remark 2.2. Applying Cauchy-Schwarz inequality and Remark 2.1, for $n\geq\beta^2+(\alpha+r)^2+3r+2\alpha+2$, we have

$$S_{n,r}^{(\alpha,\beta)}(|t-x|,x)\leq\left[\mu_{n,2,r}^{(\alpha,\beta)}(x)\right]^{\frac{1}{2}}\leq\frac{x+1}{\sqrt{n+\beta}}.$$

Lemma 2.2. *Suppose that $x\in(0,\infty)$, then for $n\geq r^2+3r+2$, we have*

$$\lambda_{n,r}(x,y)=n\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^y s_{n,k+r}(t)dt\leq\frac{(x+1)^2}{n(x-y)^2},\qquad 0\leq y<x,$$
$$1-\lambda_{n,r}(x,z)=n\sum_{k=0}^{\infty}s_{n,k}(x)\int_z^{\infty}s_{n,k+r}(t)dt\leq\frac{(x+1)^2}{n(z-x)^2},\qquad x<z<\infty.$$

Proof. The result follows directly from Remark 2.1 in the case $\alpha=\beta=0$, as for the first inequality, we have

$$\lambda_{n,r}(x,y)=n\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^y s_{n,k+r}(t)dt=\frac{S_{n,r}^{(0,0)}\left((t-x)^2,x\right)}{(y-x)^2}\leq\frac{(x+1)^2}{n(x-y)^2}.$$

Similarly, we can prove the second inequality. □

Lemma 2.3. *Let f be s times differentiable on $[0,\infty)$ such that $f^{(s-1)}(t)=\mathcal{O}(t^q)$, as $t\to\infty$ where q is a positive integer. Then for any $r,s\in N^0$ and $n>\max\{q,r+s+1\}$, we have*

$$D^sS_{n,r}^{(\alpha,\beta)}(f,x)=\left(\frac{n}{n+\beta}\right)^sS_{n,r+s}^{(\alpha,\beta)}(D^sf,x),\quad D\equiv\frac{d}{dx}.$$

Proof. First, by simple computation, we have

$$D[s_{n,k}(x)]=n[s_{n,k-1}(x)-s_{n,k}(x)].\tag{2.2}$$

The identity (2.2) is true even for the case $k=0$, as we observe that for $r<0$, $s_{n,r}(x)=0$. We shall prove the result by using the principle of mathematical induction. Using (2.2), we have

$$\begin{aligned}D\big[S_{n,r}^{(\alpha,\beta)}(f,x)\big]=&n\sum_{k=0}^{\infty}Ds_{n,k}(x)\int_0^{\infty}s_{n,k+r}(t)f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&n\sum_{k=0}^{\infty}n\big[s_{n,k-1}(x)-s_{n,k}(x)\big]\int_0^{\infty}s_{n,k+r}(t)f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&n^2\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}\big[s_{n,k+r+1}(t)-s_{n,k+r}(t)\big]f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt.\end{aligned}$$

Using (2.2), and integrating by parts we have

$$\begin{aligned}DS_{n,r}^{(\alpha,\beta)}(f,x)=&n^2\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}-\frac{D[s_{n,k+r+1}(t)]}{n}f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&\frac{n^2}{n+\beta}\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}s_{n,k+r+1}(t)f^{(1)}\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&\frac{n}{n+\beta}S_{n,r+1}^{(\alpha,\beta)}(Df,x),\end{aligned}$$

which means that the identity is satisfied for $s=1$. Let us suppose that the result holds for $s=l$ i.e.,

$$\begin{aligned}D^lS_{n,r}^{(\alpha,\beta)}(f,x)=&\Big(\frac{n}{n+\beta}\Big)^lS_{n,r+l}^{(\alpha,\beta)}(D^lf,x)\\=&n\Big(\frac{n}{n+\beta}\Big)^l\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}s_{n,k+r+l}(t)D^lf\Big(\frac{nt+\alpha}{n+\beta}\Big)dt.\end{aligned}$$

Now,

$$\begin{aligned}D^{l+1}S_{n,r}^{(\alpha,\beta)}(f,x)=&n\Big(\frac{n}{n+\beta}\Big)^l\sum_{k=0}^{\infty}Ds_{n,k}(x)\int_0^{\infty}s_{n,k+r+l}(t)D^lf\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&n\Big(\frac{n}{n+\beta}\Big)^l\sum_{k=0}^{\infty}n\big[s_{n,k+r+l-1}(x)-s_{n,k+r+l}(x)\big]\int_0^{\infty}s_{n,k+r+l}(t)D^lf\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&n^2\Big(\frac{n}{n+\beta}\Big)^l\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}\big[s_{n,k+r+l+1}(t)-s_{n,k+r+l}(t)\big]D^lf\Big(\frac{nt+\alpha}{n+\beta}\Big)dt\\=&n^2\Big(\frac{n}{n+\beta}\Big)^l\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}-\frac{D[s_{n,k+r+l+1}(t)]}{n}D^lf\Big(\frac{nt+\alpha}{n+\beta}\Big)dt.\end{aligned}$$

Integrating by parts for the last integral, we get

$$D^{l+1}S_{n,r}^{(\alpha,\beta)}(f,x)=n\Big(\frac{n}{n+\beta}\Big)^{l+1}\sum_{k=0}^{\infty}s_{n,k}(x)\int_0^{\infty}s_{n,k+r+l+1}(t)D^{l+1}f\Big(\frac{nt+\alpha}{n+\beta}\Big)dt.$$

Therefore,

$$D^{l+1}S_{n,r}^{(\alpha,\beta)}(f,x)=\Big(\frac{n}{n+\beta}\Big)^{l+1}S_{n,r+l+1}^{(\alpha,\beta)}\big(D^{l+1}f(x)\big).$$

Thus the result is true for

$$s=l+1,$$

hence by mathematical induction, the lemma is valid. □

3 Rate of converence

The class of absolutely continuous functions f defined on $(0,\infty)$ is defined by $B_q(0,\infty)$, $q>0$ and satisfying:

(i) $|f(t)|\le C_1t^q$, $C_1>0$,

(ii) having a derivative f' on the interval $(0,\infty)$ which coincides a.e. with a function of bounded variation on every finite sub-interval of $(0,\infty)$. It can be observed that for all functions $f\in B_q(0,\infty)$ possess for each $c>0$ the representation

$$f(x)=f(c)+\int_c^x\psi(t)dt,\quad x\ge c.$$

Theorem 3.1. *Let $f\in B_q(0,\infty)$, $q>0$ and $x\in(0,\infty)$. Then for n sufficiently large, we have*

$$\begin{aligned}&\Big|S_{n,r}^{(\alpha,\beta)}(f,x)-f(x)\Big|\\ &\le\frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_{x-x/k}^{x+x/k}((f')_x)+\frac{x}{\sqrt{n}}\bigvee_{x-x/\sqrt{n}}^{x+x/\sqrt{n}}((f')_x)+\frac{(1+1/x)^2}{n}(|f(2x)-f(x)-xf^{'}(x^+)|\\ &+|f(x)|)+\mathcal{O}(n^{-q})+|f'(x^+)|\frac{(x+1)^2}{n}+\frac{1}{2}\frac{x+1}{\sqrt{n+\beta}}|f'(x^+)-f'(x^-)|\\ &+\frac{\alpha+r+1-\beta x}{2(n+\beta)}|f'(x^+)+f'(x^-)|,\end{aligned}$$

where $\bigvee_a^b f(x)$ denotes the total variation of f_x on $[a,b]$, and the auxiliary function f_x is defined by

$$f_x(t)=\begin{cases}f(t)-f(x^-), & 0\le t<x,\\ 0, & t=x,\\ f(t)-f(x^+), & x<t<\infty.\end{cases}$$

Proof. Using the identity

$$\begin{aligned}f'(u)=(f')_x(u)+\frac{f'(x^+)+f'(x^-)}{2}+\frac{f'(x^+)-f'(x^-)}{2}\operatorname{sgn}(u-x)\\ +\Big[f'(x)-\frac{f'(x^+)+f'(x^-)}{2}\Big]\chi_x(u),\end{aligned}\tag{3.1}$$

where

$$\chi_x(u)=\begin{cases} 1, & u=x, \\ 0, & u\neq x. \end{cases}$$

Applying the mean value theorem, we get

$$S_{n,r}^{(\alpha,\beta)}(f,x)-f(x)=S_{n,r}^{(\alpha,\beta)}\Big(\int_x^t f'(u)du,x\Big). \tag{3.2}$$

Now, by using the above identity (3.1) in (3.2) and the fact that

$$S_{n,r}^{\alpha,\beta}\Big(\int_x^t \chi_x(u)du,x\Big)=0,$$

after simple computation, we have

$$\begin{aligned}\Big|S_{n,r}^{\alpha,\beta}(f,x)-f(x)\Big| \leq & \Big|\int_x^\infty \Big(\int_x^t (f')_x(u)du\Big) n\sum_{k=0}^{\infty} s_{n,k}(x)s_{n,k+r}(t)dt \\ & +\int_0^x \Big(\int_x^t (f')_x(u)du\Big) n\sum_{k=0}^{\infty} s_{n,k}(x)s_{n,k+r}(t)\mathrm{d}t\Big| \\ & +\frac{\Big|f'(x^+)+f'(x^-)\Big|}{2}\mu_{n,1,r}^{(\alpha,\beta)}(x)+\frac{\Big|f'(x^+)-f'(x^-)\Big|}{2}[\mu_{n,2,r}^{(\alpha,\beta)}(x)]^{1/2} \\ =&|A_{n,r}(f,x)+B_{n,r}(f,x)|+\frac{\Big|f'(x^+)+f'(x^-)\Big|}{2}\mu_{n,1,r}^{(\alpha,\beta)}(x) \\ & +\frac{\Big|f'(x^+)-f'(x^-)\Big|}{2}[\mu_{n,2,r}^{(\alpha,\beta)}(x)]^{1/2}. \end{aligned} \tag{3.3}$$

Applying Remark 2.1 and Remark 2.2 to (3.3), we have

$$\begin{aligned}\Big|S_{n,r}^{\alpha,\beta}(f,x)-f(x)\Big| \leq & |A_{n,r}(f,x)|+|B_{n,r}(f,x)|+\frac{\big|f'(x^+)-f'(x^-)\big|}{2}\frac{x+1}{\sqrt{n+\beta}} \\ & +\frac{\big|f'(x^+)+f'(x^-)\big|}{2}\frac{\alpha+r+1-\beta x}{(n+\beta)}. \end{aligned} \tag{3.4}$$

The estimation of the terms $A_{n,r}(f,x)$ and $B_{n,r}(f,x)$ will lead to proof of the theorem.

First,

$$\begin{aligned}|A_{n,r}(f,x)| &= \Big|\int_x^\infty \Big(\int_x^t (f')_x(u)du\Big) n\sum_{k=0}^\infty s_{n,k}(x)s_{n,k+r}(t)dt\Big| \\ &= \Big|\int_{2x}^\infty \Big(\int_x^t (f')_x(u)du\Big) n\sum_{k=0}^\infty s_{n,k}(x)s_{n,k+r}(t)dt \\ &\quad + \int_x^{2x} \Big(\int_x^t (f')_x(u)du\Big) d_t(1-\lambda_{n,r}(x,t))\Big| \\ &\le \Big|n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty (f(t)-f(x))s_{n,k+r}(t)dt\Big| \\ &\quad + |f'(x^+)|\Big|n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)(t-x)dt\Big| \\ &\quad + \Big|\int_x^{2x}(f')_x(u)du\Big||1-\lambda_{n,r}(x,2x)| + \int_x^{2x}|(f')_x(t)||1-\lambda_{n,r}(x,t)|dt.\end{aligned}$$

Applying Remark 2.1 with $\alpha=\beta=0$, we have

$$\begin{aligned}&|A_{n,r}(f,x)| \\ &\le n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)C_1t^{2q}dt + \frac{|f(x)|}{x^2}n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)(t-x)^2dt \\ &\quad + |f'(x^+)|\int_{2x}^\infty n\sum_{k=0}^\infty s_{n,k}(x)s_{n,k+r}(t)|t-x|dt + \frac{(1+1/x)^2}{n}|f(2x)-f(x)-xf'(x^+)| \\ &\quad + \frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_x^{x+\frac{x}{k}}((f')_x) + \frac{x}{\sqrt{n}}\bigvee_x^{x+\frac{x}{\sqrt{n}}}((f')_x). \qquad (3.5)\end{aligned}$$

To estimate the integral $n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)C_1t^{2q}dt$ in (3.5) above, we proceed as follows:

Obviously $t\ge 2x$ implies that $t\le 2(t-x)$ and it follows from Lemma 2.1, that

$$\begin{aligned}n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)t^{2q}dt &\le 2^{2q}\sum_{k=0}^\infty s_{n,k}(x)\int_0^\infty s_{n,k+r}(t)(t-x)^{2q}dt \\ &= 2^{2q}\mu_{n,2q,r}^{(\alpha,\beta)}(x) = \mathcal{O}(n^{-q}), \quad n\to\infty.\end{aligned}$$

Applying Schwarz inequality and Remark 2.1 ($\alpha=\beta=0$), the third term in right hand side of (3.5) is estimated as follows:

$$\begin{aligned}&|f'(x^+)|n\sum_{k=0}^\infty s_{n,k}(x)\int_{2x}^\infty s_{n,k+r}(t)|t-x|dt \\ &\le \frac{|f'(x^+)|}{x}n\sum_{k=0}^\infty s_{n,k}(x)\int_0^\infty s_{n,k+r}(t)(t-x)^2dt = |f'(x^+)|\frac{(x+1)^2}{nx}.\end{aligned}$$

Thus by Lemma 2.1 and Remark 2.1 ($\alpha=\beta=0$), we have

$$\begin{aligned}|A_{n,r}(f,x)|\leq&\mathcal{O}(n^{-q})+|f'(x^+)|\cdot\frac{(x+1)^2}{nx}\\&+\frac{(1+1/x)^2}{n}(|f(2x)-f(x)-xf'(x^+)|+|f(x)|)\\&+\frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_{x}^{x+\frac{x}{k}}((f')_x)+\frac{x}{\sqrt{n}}\bigvee_{x}^{x+\frac{x}{\sqrt{n}}}((f')_x).\end{aligned}\tag{3.6}$$

Applying, Lemma 2.2 with $y=x-x/\sqrt{n}$, and integrating by parts, we have

$$\begin{aligned}|B_{n,r}(f,x)|=&\Big|\int_0^x\int_x^t(f')_x(u)du\mathrm{d}_t(\lambda_{n,r}(x,t))\Big|\\=&\int_0^x\lambda_{n,r}(x,t)(f')_x(t)dt\leq\Big(\int_0^y+\int_y^x\Big)|(f')_x(t)||\lambda_{n,r}(x,t)|dt\\\leq&\frac{(x+1)^2}{n}\int_0^y\bigvee_t^x((f')_x)\frac{1}{(x-t)^2}dt+\int_y^x\bigvee_t^x((f')_x)dt\\\leq&\frac{(x+1)^2}{n}\int_0^y\bigvee_t^x((f')_x)\frac{1}{(x-t)^2}dt+\frac{x}{\sqrt{n}}\bigvee_{x-\frac{x}{\sqrt{n}}}^x((f')_x).\end{aligned}$$

Let $u=x/(x-t)$, then we have

$$\begin{aligned}\frac{(x+1)^2}{n}\int_0^y\bigvee_t^x((f')_x)\frac{1}{(x-t)^2}dt=&\frac{(x+1)^2}{n}\int_1^{\sqrt{n}}\bigvee_{x-\frac{x}{u}}^x((f')_x)du\\\leq&\frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_{x-\frac{x}{k}}^x((f')_x).\end{aligned}$$

Thus

$$|B_{n,r}(f,x)|\leq\frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_{x-\frac{x}{k}}^x((f')_x)+\frac{x}{\sqrt{n}}\bigvee_{x-\frac{x}{\sqrt{n}}}^x((f')_x).\tag{3.7}$$

The required result is obtained by combining (3.4), (3.6) with (3.7). □

As a consequence of Lemma 2.3, we have the following corollary:

Corollary 3.1. Let $f^{(s)}\in DB_q(0,\infty)$, $q>0$ and $x\in(0,\infty)$. Then for n sufficiently large, we

have

$$\begin{aligned}
&\left|D^s S_{n,r}^{(\alpha,\beta)}(f,x)-f^{(s)}(x)\right| \\
\leq&\frac{(x+1)^2}{nx}\sum_{k=1}^{[\sqrt{n}]}\bigvee_{x-x/k}^{x+x/k}((D^{s+1}f)_x)+\frac{x}{\sqrt{n}}\bigvee_{x-x/\sqrt{n}}^{x+x/\sqrt{n}}((D^{s+1}f)_x) \\
&+\frac{(1+1/x)^2}{n}(|D^s f(2x)-D^s f(x)-xD^{s+1}f(x^+)|+|D^s f(x)|)+\mathcal{O}(n^{-q}) \\
&+\frac{(x+1)^2}{nx}|D^{s+1}f(x^+)|+\frac{1}{2}\frac{x+1}{\sqrt{n+\beta}}|D^{s+1}f(x^+)-D^{s+1}f(x^-)| \\
&+\frac{1}{2}|D^{s+1}f(x^+)+D^{s+1}f(x^-)|\frac{\alpha+r+1-\beta x}{n+\beta},
\end{aligned}$$

where $\bigvee_a^b f(x)$ denotes the total variation of f_x on $[a,b]$, and f_x is defined by

$$D^{s+1}f_x(t)=\begin{cases} D^{s+1}f(t)-D^{s+1}f(x^-), & 0\leq t<x, \\ 0, & t=x, \\ D^{s+1}f(t)-D^{s+1}f(x^+), & x<t<\infty. \end{cases}$$

4 Asymptotic formula

We consider the class $L[0,\infty)$ of all measurable functions defined on $[0,\infty)$ such that

$$L[0,\infty):=\left\{f:\int_0^\infty e^{-nt}f(t)dt<\infty \text{ for some positive integer } n\right\}.$$

It can be observed that this class is bigger than that of all integrable functions on $[0,\infty)$.

Further we consider

$$L_\alpha[0,\infty):=\left\{f\in L[0,\infty):f(t)=\mathcal{O}(e^{\alpha t}),\ t\to\infty,\ \alpha>0\right\}.$$

We have the following asymptotic formula by using Lemma 2.1.

Theorem 4.1. *Let $f\in L_\alpha[0,\infty)$ and suppose it is bounded on every finite subinterval of $[0,\infty)$ having a derivative of order $r+2$ at a point $x\in(0,\infty)$, then we have*

$$\lim_{n\to\infty} n[(S_{n,r}^{(\alpha,\beta)})^{(r)}(f,x)-f^{(r)}(x)]=(\alpha+r+1-\beta x)f^{(r+1)}(x)+xf^{(r+2)}(x).$$

The proof follows along the line of [1].

Acknowledgements

The authors are extremely thankful to the referee for valuable comments, leading to the overall improvements in the article.

References

[1] V. Gupta, V., Simultaneous approximation by Szász-Durrmeyer operators, The Math. Student, 64(1-4) (1995), 27–36.
[2] M. K. Gupta, M. S. Beniwal and P. Goel, Rate of convergence for Szász-Mirakyan-Durrmeyer operators with derivatives of bounded variation, Appl. Math. Comput., 199(2) (2008), 828–832.
[3] S. M. Mazhar and V. Totik, Approximation by modified Szász operators, Acta Sci. Math., 49 (1985), 257–269.
[4] H. M. Srivastava and X. M. Zeng, Approximation by means of the Szász-Bézier integral operators, Int. J. Pure Appl. Math., 14(3) (2004), 283–294.
[5] D. D. Stancu, Approximation of function by a new class of polynomial operators, Rev. Roum. Math. Pures et Appl., 13(8) (1968), 1173–1194.

J Nanopart Res (2013) 15:1697
DOI 10.1007/s11051-013-1697-z

RESEARCH PAPER

Synthesis of CdS nanocrystals in poly(3-hexylthiophene) polymer matrix: optical and structural studies

Vikash Agrawal · Kiran Jain · Leena Arora · S. Chand

Received: 24 October 2012 / Accepted: 3 May 2013

Abstract CdS nanocrystals (NCs) were directly synthesized in P3HT matrix by decomposition of single-molecule precursor compound. In this process, a molecular precursor solution was mixed with the polymeric solution. On heating the solution to the decomposition temperature of the precursor compound, NCs were formed in situ at temperatures as low as 120 °C. The effects of the precursor concentration on the optical properties of the composite were studied. The results showed evidence of charge transfer and size variation depending on NCs concentration. CdS phase can be formed using this process at 120 °C temperature as was evident from the X-ray diffraction studies. Transmission electron microscope results confirm formation of monodispersed CdS nanoparticles of average size 4 nm. A possible mechanism of the CdS film formation was also investigated. UV–Vis measurements show that these CdS composites possess a direct band gap energy higher than 2.45 eV depending on the concentration of P3HT, indicating that the nano size can be controlled by the concentration of polymer additive in the composite. A higher concentration of P3HT showed more blue shift.

Keywords Nanocrystals · Polymers · Precursor · Ligands · Composites

V. Agrawal
CSIR-SRF, Organic and Hybrid Solar Cell Group, National Physical Laboratory, Dr. K. S. Krishnan Road, New Delhi 110 012, India

K. Jain (✉) · S. Chand
Organic and Hybrid Solar Cell Group, National Physical Laboratory, Dr. K. S. Krishnan Road, New Delhi 110 012, India
e-mail: kiran@mail.nplindia.org

L. Arora
Department of Physics, Delhi Technological University, New Delhi 110042, India

Introduction

The development of photovoltaic (PV) technologies has attracted increasing attention due to the world-wide increasing energy demand as well as the need for replacing depleting energy resources such as fossil fuels by renewable energy resources. Polymer solar cell is one of the most promising technologies due to its compatibility with solution-based roll-to-roll manufacturing as well as it has thin, flexible, and lightweight features (Hoppe and Sariciftci 2004; Wang et al. 2012). Efficiency reaching a level of over 10 % has already been demonstrated (Mayukh et al. 2012). Recently, new PVs based on solution-processed inorganic nanocrystals (NCs) have attracted considerable attention, and PV devices based on various NCs have been demonstrated (Habas et al. 2010). Hybrid organic–inorganic composites are potential PV device

Published online: 23 May 2013

systems because the desirable characteristics of both organic and inorganic components can be included within a single composite material (Reiss et al. 2010; Saunders and Turner 2008). Such composites show advantages of tunability of photophysical properties of the quantum dots and they also retain the polymer properties like solution processing, fabrication of devices on large and flexible substrates. Due to different electron affinities of QD and polymer, a built-in potential is generated at polymer–QD interface, which assists in the charge transfer between polymer and QD. Various composite systems comprising conjugated polymers, serving as hole conductors in conjugation with NCs of CdSe, CdTe, TiO_2, ZnO, PbS, CdS have been reported (Huynh et al. 2002; Dayal et al. 2010; Jeltsch et al. 2012; Verma et al. 2009; Liu et al. 2008; Beek et al. 2006; Wang et al. 2008; Han et al. 2011). Up to now, the organic–inorganic hybrid thin film solar cells have shown power-conversion efficiency of $\sim$3–5 %. The stability of most polymer–nanocomposite devices exceeds that of fully organic-based devices dramatically, making them especially attractive for application purposes.

The nanocrystal/conducting polymer composites, used for such opto-electronic applications are produced either by synthesizing the NCs separately and then blending them with a conducting polymer or by synthesizing the NCs directly in the conducting polymer. In the first process, NCs are required to be homogeneously dispersed in the polymer matrix, to constitute an interpenetrating network in order to obtain efficient exciton separation and charge transport in the active layer. NCs are generally synthesized using organic capping ligands to stabilize their size and to prevent aggregation (Talapin et al. 2010). Although the incorporation of the surface ligands into an organic/inorganic composite can improve distribution of NCs in a conjugated polymer, the insulating nature of such organic ligands results in poor electronic communication among the NCs (Moule et al. 2012). The initial ligands are usually exchanged for shorter ones such as pyridine, *n*-butyl amine, *t*-butylthiol to promote polymer–NC charge transport (Zhou et al. 2010; Greaney et al. 2012).

These problems can be bypassed by directly growing NCs in a conducting polymer solution, without surfactants or ligands. In situ synthetic method was employed to synthesize PbS NCs/MEH–PPV composite and it was shown that the polymer chains can act sterically to stabilize nanocrystal growth in solution (Stavrinadis et al. 2008; Warner and Watt 2006). Synthesizing the NCs directly in a conducting polymer eliminates the need for mixing or blending and represents the simplest technique for forming nanocomposite materials, suitable for opto-electronic applications. Therefore, development of semiconductor NCs, without the use of surfactants and/or ligands, becomes more technically interesting as it can provide improved power-conversion efficiency for a quantum-dot-based PV device. Up to now, various inorganic NCs have been synthesized in situ, in presence of polymers (Stavrinadis et al. 2008; Warner and Watt 2006; Coates et al. 2010; Dayal et al. 2009; Khan et al. 2011; Maier et al. 2011). Direct synthesis of CdSe QDs was reported in 1-octadecene solution in presence of P3HT polymer (Dayal et al. 2009). A similar route was used to synthesize CdTe NCs in P3HT matrix (Khan et al. 2011). These synthetic procedures yield homogeneous dispersion of NCs in polymer; however, in these techniques after the synthesis of the NCs, the composite NCs need to be separated out from the solution, purified to remove trioctylphosphine (TOP) or other such additives and re-dispersed back in a different solvent for making a thin film. It is not easy to control the stoichiometry of the semiconductor compound using this technique. This drawback can be overcome by using single-molecule precursor to prepare NC, wherein both the metal and the nonmetal parts can be included in the same compound and subsequent precursor decomposition is used to NC nucleation and growth. Thermal decomposition (thermolysis) of single-molecule precursors (SMP) is an efficient route for synthesizing chalcogenides (Nair et al. 2005; Plante et al. 2010; Shen et al. 2011). As an example, the synthesis of PbS NCs, a molecular precursor such as xanthate or carbamate is heated, this single molecule contains both the metal and chalcogen atoms, and on decomposition PbS NCs are nucleated together with the release of some gases, which are then grown to the desired size by controlled reaction of precursor molecule. These precursors are non-air sensitive, stable for a period of months, easy to synthesize and pyrolyse cleanly to give the highest yield. Some people used such process to synthesize PbS NCs in polystyrene (Nair et al. 2005).

CdS is one of the promising photoactive materials because it is a good electron acceptor, and has large band gap (2.5 eV). A large variety of methods were

reported to synthesize CdS polymer composites (Pradhan et al. 2003; Laera et al. 2011; Masala et al. 2011; Borriello et al. 2011; Di Luccio et al. 2006; Laera et al. 2011; Liao et al. 2009, 2010, 2011; Ren et al. 2011). CdS nanorods have grown in situ in a P3HT template, the sulfur atom of the thiophene ring giving an anchoring site for the nucleation of the NCs. One of the common SMP, employed for the generation of metal sulfide is *N*,*N*-dialkyl dithiocarbamate. This method of preparation involves basically injecting the precursor into hot coordination solvents such as trioctylphosphine oxide (TOPO) under nitrogen atmosphere. High temperature, in the range of 200–350 °C, is thus required, and stringent experimental conditions, as well as inert atmosphere are normally needed. Very recently, metal salts of alkylxanthate were used to produce CdS NCs capped with HDA at relatively milder conditions at 50–150 °C (Nair et al. 2002). It was found that Lewis base such as HDA, TOP or tributylphosphine could lower the reaction temperature for the alkylxanthate precursors. CdS NCs were synthesized by thermolysis of a cadmium bis thiolate precursor, without the use of any surfactant or other stabilizing agent (Resta et al. 2010). $Zn_xCd_{1-x}S$ alloyed NCs were synthesized at about 200 °C, by thermolyzing a mixture of cadmium ethyl xanthate (CdX) and zinc ethyl xanthate precursors in colloidal solvents (Li et al. 2005). CdS NCs capped with TOPO have been synthesized by a single source route using CdX as a precursor (Nair et al. 2002). When the precursor is heated well above the decomposition temperature of the precursor, sudden decomposition takes place, which results in a burst of homogeneous nucleation. Further heating controls the growth. As these results indicate, the use of single source precursor is of great advantage in synthesizing good quality of NCs, as they have to be heated to only moderately high temperatures. Recently, Leventis et al. (2010), Dowland et al. (2011) have reported a method to deposit P3HT–CdS composite directly on the substrate. Precursor cadmium xanthate and pyridine react on heating to form CdS and form gaseous side products. The casted P3HT:precursor solid films were heated to 120–160 °C to allow the precursor decomposition, and the CdS/NCs are formed within the polymer matrix.

In the present work, the synthesis of CdS NCs was carried out using single-molecule precursor, CdX. Here, Cd and sulfur both are present in the single molecule used as precursor, which on heating/pyrolysis produce CdS, therefore stoichiometry can be controlled easily. The present investigation reports the synthesis of CdS NCs in P3HT matrix with different concentrations of P3HT and CdS (1:1, 1:2, 1:4 and 1:8 weight%). These P3HT–CdS nanocomposites are synthesized in common solvent di-chlorobenzene (DCB), from which thin films can be readily cast. The size, optical and structural properties of the resulting composite has been investigated. The basic aim of the work is to understand the effect of polymer concentration on the growth of NCs. Structural and morphological studies revealed that hexagonal CdS NCs have been successfully synthesized in P3HT matrix. Optical measurements of nanocomposite films show that photoinduced charge separation occurs at the P3HT–CdS interfaces, indicating it as a promising approach for the fabrication of efficient organic–inorganic hybrid PV devices.

Experimental section

Synthesis of cadmium ethyl xanthate

Cadmium chloride (0.005 mol), dissolved in 20 ml distilled water was allowed to react with a solution of potassium ethyl xanthate (0.01 mol) in 20 ml distilled water. The pale yellow precipitate (CdX) formed was filtered, washed with distilled water and dried at 40–50 °C.

Synthesis of CdS/polymer composite

Regioregular poly(3-hexylthiophene-2,5-diyl) P3HT with average mol. wt. 54,000–75,000, electronic grade 99.995 % from Sigma Aldrich was used to prepare the composites. In a typical synthesis of P3HT:CdS = 1:1, 15 mg of CdX was kept in a culture tube (xanthate compound was taken 3 times the weight of P3HT polymer to compensate the weight loss of metal xanthate during decomposition to yield CdS) and a few drops of pyridine were added and mixed. 5 ml of *o*-dichlorobenzene solvent was added to dissolve the above compound well and 5 mg of P3HT polymer was added. The addition of pyridine causes the xanthate to dissolve well in DCB. The mixture was stirred for

10 min, after which the precursors had dissolved and the color of the reaction solution changed from light-yellow to dark-yellow orange. The precursor–polymer solution was then either heated directly under nitrogen flow for 30 min at 120 °C or casted on a substrate to make a film and then heated. After the completion of the reaction, the reaction product was characterized directly in solution for PL measurements, while for X-ray diffraction (XRD) films were casted from synthesized composite NCs. Similarly, other compositions of P3HT containing different weight ratios of CdS were synthesized and are designated as 1:1, 1:2, 1:4 and 1:8, for 15, 30, 60 and 120 mg of CdX precursor in 5 mg P3HT and 5 ml DCB. Another batch was prepared under the same conditions for CdS nanocrystal synthesis (using 60 mg CdX) without adding the polymer, for comparison.

Characterization techniques

The synthesized P3HT–CdS polymer composites were characterized using XRD, TEM, SEM, HRTEM, UV absorption, photoluminescence and time-resolved florescence spectroscopy. The XRD studies were carried out using Rigaku Miniflex-II diffractometer (Cu K$\alpha = 1.54$ Å). Transmission electron microscope (TEM) and scanning electron microscope (SEM) studies were made using a JEM-2100F (JEOL electron microscope) 200 kV and Zeiss EV MA-10 for microstructural and surface morphology studies. A drop of dispersed solution of the CdS/polymer composite was carefully transferred to a carbon-coated copper grid followed by evaporation of the solvent for TEM studies. For SEM studies, a thin gold layer was sputtered on to the film surface. Absorption spectra were recorded using UV 1800 Shimadzu UV Spectrometer. The photoluminescence measurements were carried out using excitation wavelength of 410 nm, at room temperature using a home-assembled system consisting of a two-stage monochromator, a photomultiplier tube with a lock-in amplifier for PL detection. Fourier transform infrared spectroscopy (FTIR) spectra were recorded with a single beam Perkin Elmer instrument (Spectrum BX-500). This spectrometer allows us to collect spectra in IR (mid, far and near) spectral ranges. PL decay measurements were recorded using Edinburgh Laser Spectrometer FLS 920 (wave length 480 with pulse width 93.9).

Results and discussion

Cadmium ethyl xanthate has been proven to be an efficient single-molecule precursor for the synthesis of CdS nanocrystallites. This precursor is easy to synthesize, air stable and pyrolyzes cleanly to provide environment friendly good yields. Thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC) of CdX was carried out up to 300 °C temperature. The TGA curve shown in Fig. 1 reveals that the compound decomposes at about 160 °C and DSC curve reveals that the reaction is endothermic. The addition of pyridine decreases the decomposition temperature, as CdS NCs can be formed at 120 °C after pyridine addition.

$$Cd(C_2H_5OCS_2)_2 \xrightarrow{HEAT} CdS + \text{gases}$$

The CdS NCs were synthesized in high temperature boiling point solvent DCB, which allows CdS

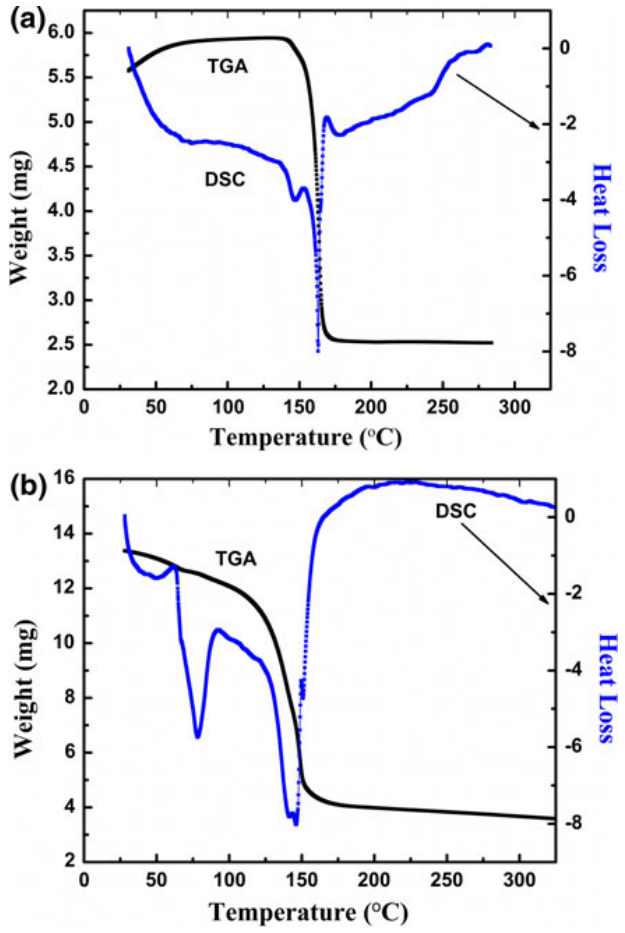


Fig. 1 **a** TGA and DSC curves for cadmium ethyl xanthate, **b** TGA and DSC curves for cadmium ethyl xanthate (with pyridine)

nucleation to occur at 120 °C. In the present protocol, nucleation starts at the decomposition temperatures of the precursor compound in one-pot synthesis, which results in highly crystalline and mono dispersed CdS NCs in the polymer matrix. The P3HT acts here as a macromolecule which stimulates a structure-directing template to manipulate the geometrical development of CdS nanostructures. In surfactant-assisted synthesis, growth of NCs is controlled by electrostatic interactions of the surfactant functional group and steric hindrance of surfactant side alkyl chains. P3HT provides a combination of both effects as it contains an electron donating sulfur functionality, a potential anchorage for the nucleation, and growth of NCs along with steric hindrance due to long hexyl side chains. Proposed mechanism for in situ growth of the CdS NCs in the P3HT matrix can be given in three stages: (a) P3HT and CdX precursors were dissolved in DCB using few drops of pyridine; (b) on heating, the CdX precursor decomposes to produce CdS together with some volatile products; (c) in situ-generated CdS NCs were capped by the P3HT polymer matrix.

TGA and DSC studies

Thermal decomposition of precursor CdX was studied by TGA and DSC, without pyridine addition, Fig. 1a as well as with pyridine addition, Fig. 1b. TGA was carried out from room temperature to 350 °C. The TGA curve in Fig 1a shows that significant weight loss of 57 % takes place in temperature region starting at 145 °C until 165 °C, because of thermal decomposition of xanthate complex into CdS. The end products are cadmium sulfide and volatile compounds. Since the decomposition occurs around 145 °C, the synthesis was carried out below this temperature only. After pyridine addition weight loss begins from 100 to 150 °C.

The DSC plot in Fig 1a reveals several peaks and dips. The sharp dips at temperature 147 and 162 °C represent that the reaction is endothermic and dip at temperature 170 °C, the reaction is exothermic. The enthalpy changes associated with the events occurring are given by the area under the sharp dips. The sharp dip at 147 °C accompanied with a significant weight loss indicates that the major decomposition of CdX takes place at this temperature. In Fig. 1b, two main dips were observed at 77 and 145 °C. The weight loss TGA curve and DSC show that after pyridine addition, decomposition of xanthate begins from 100 °C.

Spectral study (FTIR)

The formation of P3HT–CdS nanocomposites was confirmed by the FTIR spectra as shown in Fig. 2. Absorption band of P3HT at 2,960 cm^{-1} was assigned to the asymmetric CH3 stretching. The bands at about 2,920 and 2,848 cm^{-1} correspond to the C–H of CH2 or CH3 symmetric and asymmetric stretching vibration; the bands at about 1,508 cm^{-1} correspond to thiophene ring C=C stretching vibration; the bands at about 1,370 and 1,450 cm^{-1} correspond to symmetric and asymmetric deformation vibrations, respectively; the bands at about 1,199 cm^{-1} correspond to C–C single resonance absorption; 1,014 and 828 cm^{-1} correspond to C–H out-of-plane bending and the bands at 716 cm^{-1} correspond to methylene vibration. In nanocomposites of P3HT–CdS, the intensity of peaks corresponding to C–S bond and aromatic C–H out-of plane stretching decreases. Also a shift of almost all the peaks belonging to P3HT shift to lower wave number side as compared to P3HT in the composite, this confirmed the chemical interaction between P3HT and CdS, indicating the enhancement of the C–S bond energy. An extra peak appears in the composite present at 1732 cm^{-1}, which might be due to C–O or Cd–O. The characteristic vibration peaks of the Cd–S bond at 265 and 405 cm^{-1} could not be detected in this

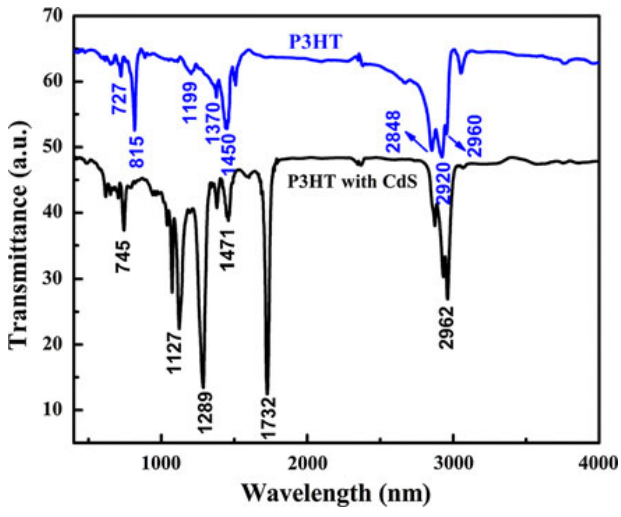


Fig. 2 FTIR spectra of P3HT and P3HT–CdS composite

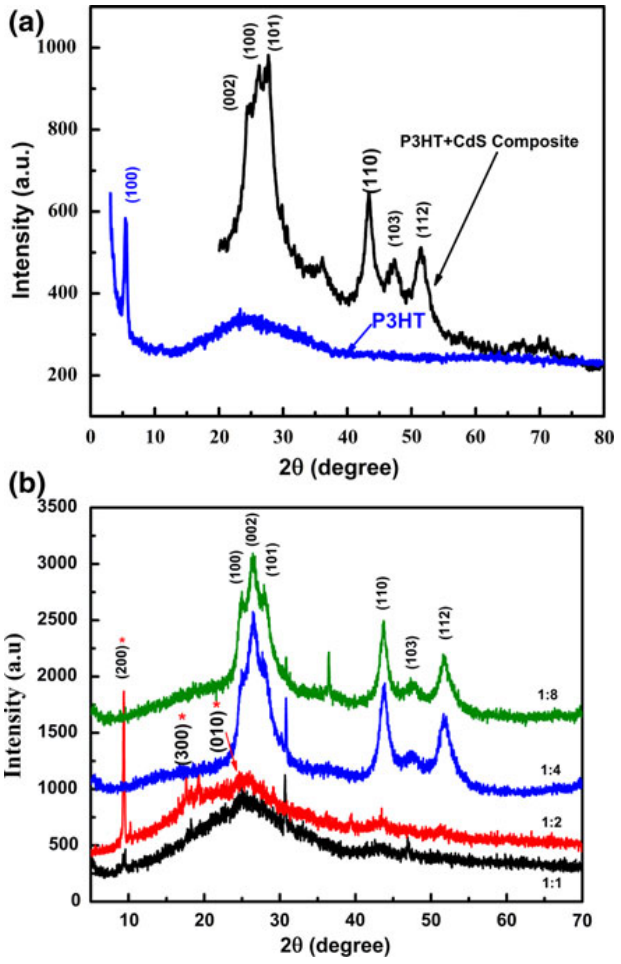


Fig. 3 XRD spectra for **a** CdS and P3HT; **b** P3HT–CdS composite for different composition ratios

experiment, since this is beyond the detection limit of our instruments.

Figure 3a shows the powder XRD derived from the pyrolysis of CdX at 120 °C. As is well known, CdS can exist in two crystalline forms, cubic and hexagonal. The hexagonal phase is thermodynamically metastable, which is usually stable at high temperatures, while cubic phase is more thermodynamically stable and forms at low temperatures. The hexagonal/wurtzite phase forms on heating CdS to high temperatures. As shown in Fig. 1a, the (110), (104) and (112) planes of wurtzite CdS are clearly distinguishable in the diffraction pattern. Figure 3a shows the pattern for pure P3HT film derived after similar heat treatment, this shows that (200) peak belongs to P3HT. Figure 3b shows the XRD pattern for P3HT–CdS with different CdS concentrations.

P3HT/CdS composites clearly show the existence of hexagonal phase of CdS at higher CdS concentration; however, for low CdS concentrations the (002) peak is broadened, and mixed with other CdS peaks. The XRD peaks are broadened in the composite as compared to those of the bulk CdS which is an indication that the crystals are in the nanosize regime formed at 120 °C. In an earlier study, wurtzite ZnS NCs are synthesized on thermolysis of zinc ethyl xanthate at 150 °C in presence of HDA and TOP (Liu 2006). In contrast, in the present study, we have used pyridine instead of HDA or TOP, and as shown in Fig. 1a, hexagonal CdS is formed at a lower temperature of 120 °C. Levantis et al. (2010), Dowland et al. (2011) have reported synthesis of CdS in polymer at 120–160 °C using a solution-processible CdX pyridine adduct. From the XRD results Fig. 3b, confirmed the in situ formation of CdS NCs. The diffraction peaks at (200), (300) and (010) can be associated with the P3HT as reported by Borreiello et al. (2011) and at 24.99°, 26.53°, 27.92°, 43.83°, 47.79°, 51.75° may be assigned to (100), (002), (101), (110), (103) and (112) planes of CdS material, respectively, which are consistent with those in the literature and JCPDS card (no. 65–3414) also confirms the formation of hexagonal phase of CdS nanocrystal in polymer nanocomposites. The intensity of crystalline peaks of P3HT decreases as we increased the ratio of nanocrystal precursor in polymer nanocrystal composite. The dimension of CdS nanocrystal in nanocomposite and CdS without composite can be calculated using Debye–Scherrer equation

$$d = 0.9\lambda/(\beta \cos\theta)$$

where β is the line width at an angle 2θ and λ is X-ray wave length.

The size of the crystallites was determined using the (110) reflection at $2\theta = 43.78^{\circ}$, since it is difficult to separate out lower angle peaks for low concentration of NCs. Using this formula we can observe that crystal size of CdS NCs in 1:1 polymer composite is about 2 nm and for polymer:CdS ratios of 1:4 and 1:8 composites, it is 5 and 6 nm, respectively. This result clearly shows that as we increase the amount of CdS in polymer composite, the nanocrystal size increases. Here the P3HT can act as a macromolecule template (surfactant) to control the geometrical development of CdS NCs. Thus, we can conclude that the presence of polymer influences the nanocrystal size in the solution. Without addition of polymer, the CdS nanocrystal synthesized in the same solvent under similar

conditions showed a nanocrystal size of 6 nm, i.e., almost near to that observed at higher CdS content (1:8 wt. ratio).

Previous studies of in situ synthesis method have shown that larger molecular weight poly[2-methoxy-5-(2′-ethylhexyloxy)-*p*-phenylene vinylene] (MEHPPV) produces smaller NCs (Warner and Watt 2006). This finding supports the concept that polymer chains act sterically to stabilize nanocrystal growth in solution. Thermal decomposition of molecular precursors is one of the most efficient pathways for synthesizing CdS NCs. For example, in the thermolysis process, a molecular precursor such as a xanthate or carbamate is heated so as to nucleate CdS NCs, which are then grown to the desired size by controlling reaction of precursor molecules, reaction time and temperature. The average size of CdS NCs is smaller than the Bohr diameter of CdS bulk exciton, this suggests that quantum confinement effect can be expected for these nanocrystals. The results presented demonstrate that CdX effectively works as a single-molecule precursor for in situ synthesis of CdS NCs and the size of NCs depended on the concentration of polymer in the composite.

TEM studies

Transmission electron microscopy (TEM) image and electron diffraction patterns of the synthesized P3HT–CdS nanocomposites in 1:4 ratio at 120 °C are shown in Fig. 4a–d. A difference in contrast at different places indicates that the CdS NCs are capped by P3HT. However, this interaction between polymer and NCs indicates that nanocomposites have potential for the charge transfer at polymer–NCs interfaces, which is further confirmed by the optical studies. The mechanism of this interaction has revealed that the sulfur atom of P3HT can interact with the CdS NCs by dipole–dipole interaction and CdS NCs have been deposited uniformly and compactly on or in between the P3HT chains to form NCs. Figure 4a shows a number of CdS NCs of almost same size, distributed randomly. Figure 4b shows the formation of small sized,

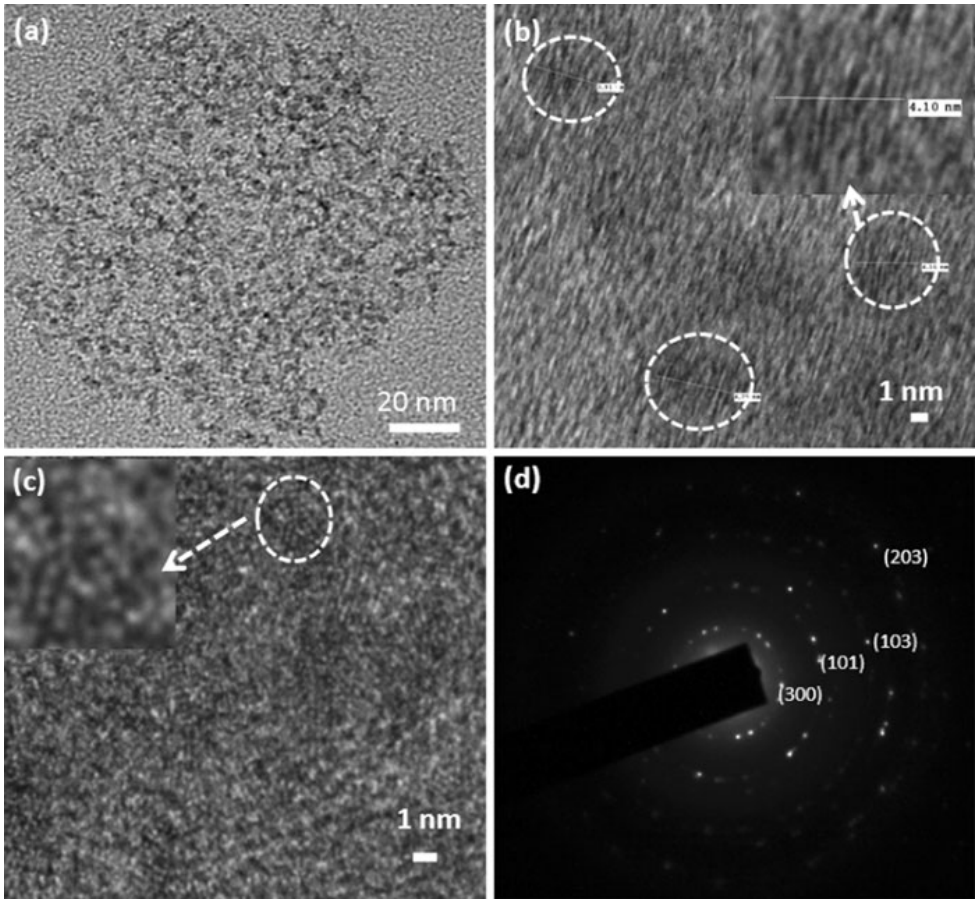
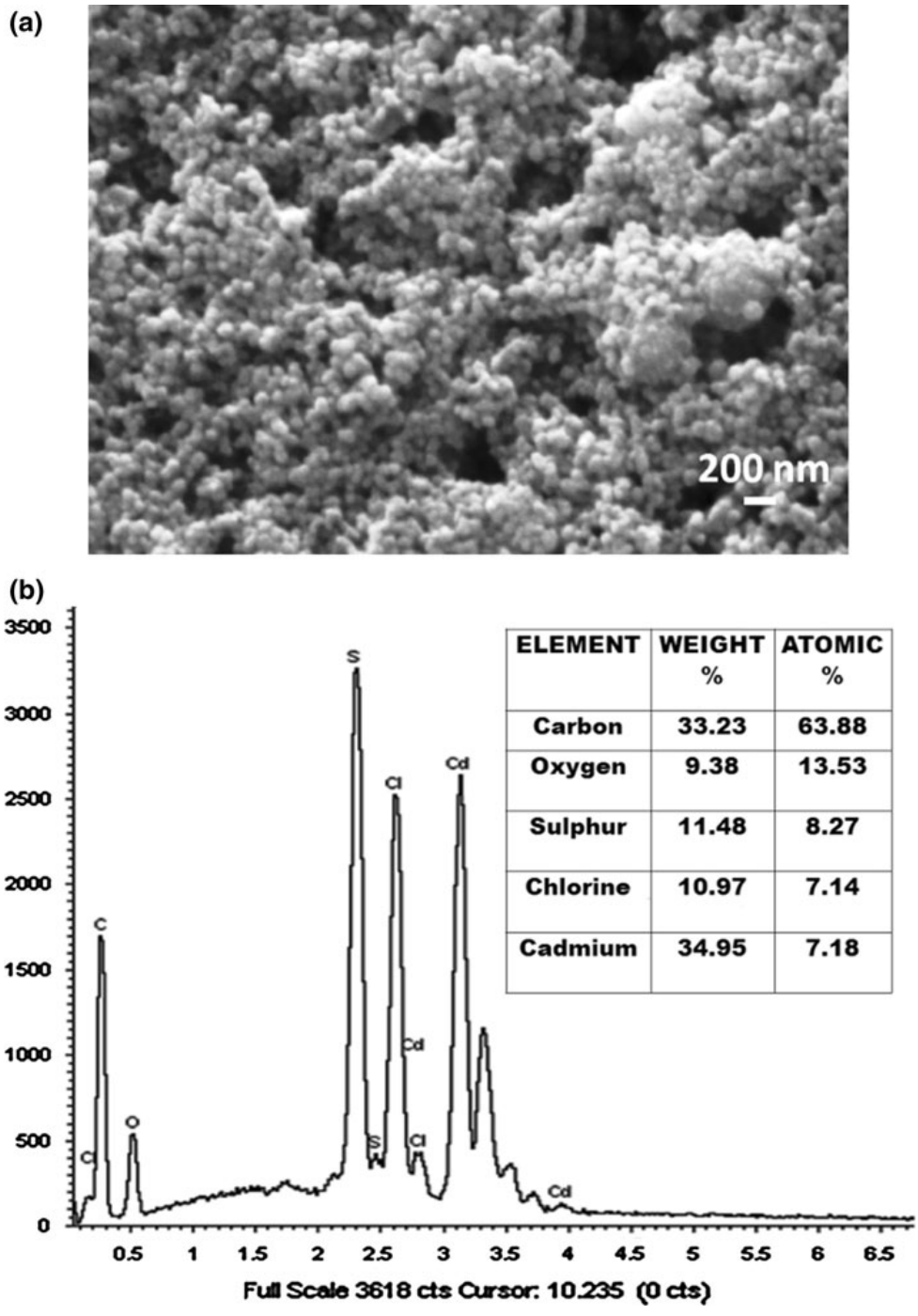


Fig. 4 **a–c** TEM images of the synthesized P3HT–CdS (1:4) at 120 °C and **d** SAED pattern of P3HT–CdS

monodispersed circled CdS NCs of average size 4 nm. Figure 4b, c show a high magnification images of corresponding NCs in the inset. The presence of sharper lattice fringes in high-resolution image shows good crystallinity of NCs and these results are in good agreement with XRD results. The selected area electron diffraction pattern, of CdS NCs, Fig. 4d, also confirms the single crystalline nature of NCs and the CdS NCs with plane (101), (103) and (203) of diffraction rings, corresponding to d spacing 3.135, 1.9 and 1.36 nm. Plane (300) of diffraction ring corresponding to d spacing 4.96 nm belongs to the polymer P3HT.

Scanning electron microscopy of polymer–CdS nanocomposites

In order to understand the morphology of the as synthesized CdS NCs in polymer matrix, SEM was



ELEMENT	WEIGHT %	ATOMIC %
Carbon	33.23	63.88
Oxygen	9.38	13.53
Sulphur	11.48	8.27
Chlorine	10.97	7.14
Cadmium	34.95	7.18

Fig. 5 SEM micrographs of P3HT–CdS composite **a** microstructure **b** EDS analysis

preformed. Figure 5 gives the typical morphologies of the P3HT–CdS films prepared at 120 °C with CdS of 60 mg, which shows the presence of large micrometer sized agglomerates and CdS NCs, uniformly distributed on the film surface. To further investigate the composition of the CdS in composite film, the energy dispersive X-ray spectroscopy (EDX) spectrum (Fig 5b) was obtained. EDX analysis of the CdS–polymer composite demonstrates that the chemical component consists of Cd, C, Cl and S and table (inset) shows approximate relative ratio of Cd:S $\approx$ 1:1. This ratio agrees with the nominal CdS stoichiometry.

The AFM images of pure P3HT and P3HT/CdS composite films on glass plate for pure P3HT, P3HT:CdS of 1:1, 1:4 and 1:8 ratio are shown in Fig. 6a–d, respectively. The composite with high NC content is more rough as compared to pure P3HT film. The surface roughness values of films increases on increasing the nanocrystal content; for the pure P3HT film it was 2 nm, but increased to 3.5 nm for 1:4 and it was 11.5 nm for 1:8 sample. In Fig. 6d, partial aggregation was detected at the surface of sample containing high amount of NCs (1:8). Contrary to organic materials, the inorganic NCs are not well dispersed in organic matrix, when mixed in large amount, the high content of CdS leads to some aggregation. The AFM image of the neat P3HT film shows closely packed polymer nodules, leading to a surface roughness of 2 nm. The images of the P3HT–CdS nanocomposite (b–d) clearly shows a uniform morphology with small nanocrystallites, aggregated to nodule size of <40 nm, thus the segregation of the NCs or the polymer in these composites is not high. These composites were synthesized after spin coating

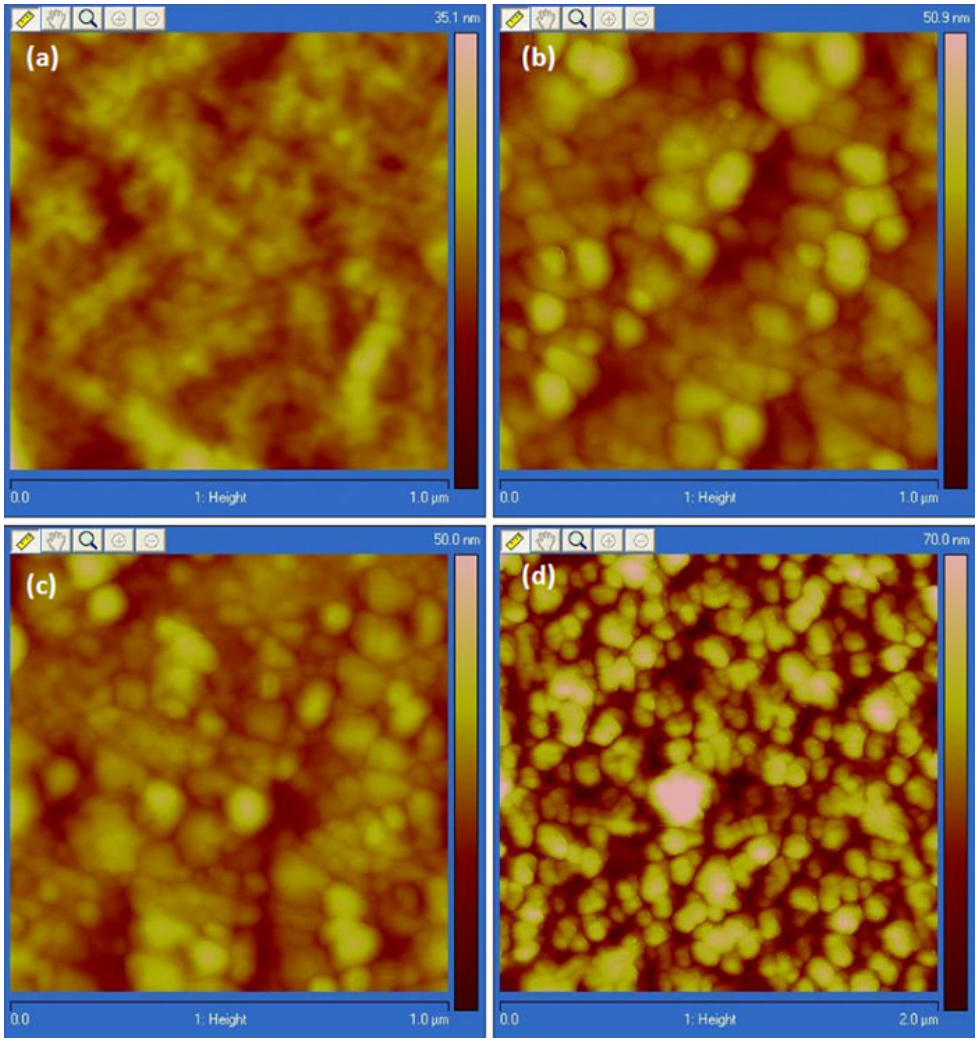


Fig. 6 AFM for **a** P3HT, **b** P3HT–CdS 1:2, **c** P3HT–CdS 1:4 **d** P3HT–CdS 1:8

the precursor–polymer solutions on glass plate and then heating to 120 °C for xanthate decomposition. The uniform layer synthesis through this route confirms that the technique is quite suitable to construct hybrid solar cell directly from the precursor–polymer solution. As pointed out earlier from XRD results, the effect of polymer on the growth of NCs for polymer–CdS 1:8 composite was negligible, and some aggregation of NCs was observed for this sample as the passivation of CdS NCs through polymer was lesser in this case.

Optical properties of CdS–P3HT composite

Figure 7a shows the UV–visible absorption spectra of polymer/CdS nanocomposite films on glass plate. On heating up to 120 °C, xanthates start to decompose and CdS nucleation gets initiated at the decomposition temperature of xanthate, and the consequent growth of CdS NCs is initiated by this process. The optical band gap of bulk CdS is 2.42 eV, and its main absorption range is almost below 550 nm. Therefore, addition of CdS NCs contributes significantly in the absorbance within 300–550 nm range. The concentration of P3HT in all solutions was constant. This allows any changes in the optical spectra to be attributed to the CdS nanocrystals. The peaks in the absorption spectra in Fig. 7, at 550 and 625 nm, are attributed to the P3HT which corresponds to the π–π^* transition of the conjugated chain in the P3HT. The P3HT–CdS samples, however, show large adsorption feature within a broad range between 400 and 600 nm, the intensity of which increases with the concentration of CdS addition, which is the characteristic absorption of the CdS material. A similar feature has been observed by Levantis et al. (2010) confirming that CdS formation occurred during heating of the samples.

Since CdS is a direct band gap semiconductor, its absorption coefficient (α) and band gap (E_g) satisfy the equation:

$$(\alpha h\upsilon)^2 = A\left(h\upsilon - E_g\right) \quad (1)$$

The band gap (E_g) is obtained by extrapolation of the plot of $(\alpha h\upsilon)^2$ versus ($h\upsilon$) as in Fig. 7b and the Tauc's plot shows that band gap of CdS in the nanocomposite varies from 2.97 to 2.41 eV, as we increase the concentration of CdS from 1:1 to 1:8. These results reveal that with increase in polymer concentration the nanocrystal size of CdS decreases as

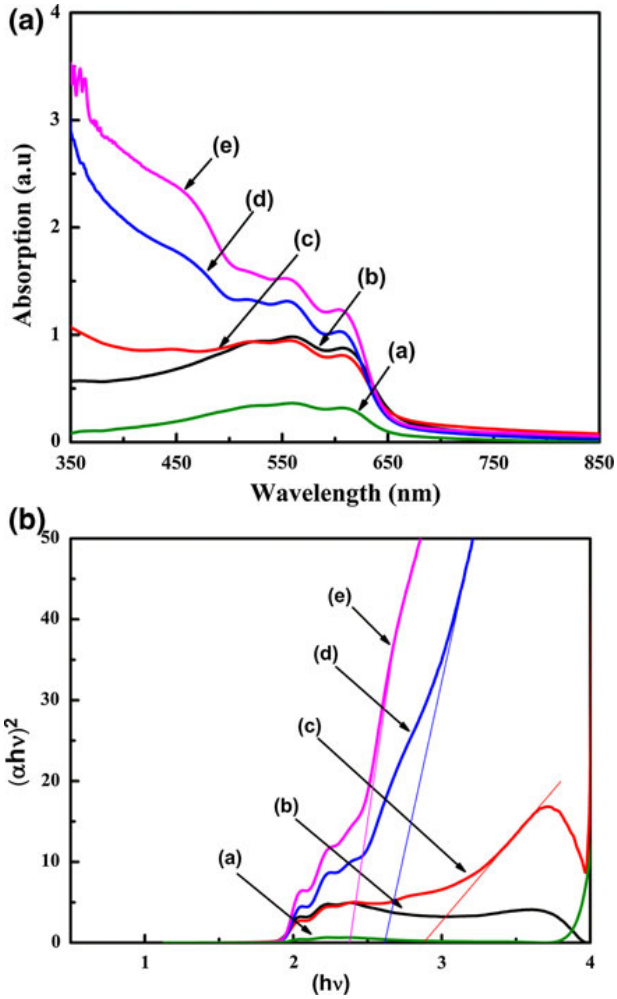


Fig. 7 **a** Absorption spectra **b** Tauc's plot for (*a*) P3HT (*b*) P3HT–CdS 1:1 (*c*) P3HT–CdS 1:2 (*d*) P3HT–CdS 1:4 (*e*) P3HT–CdS 1:8

well as the band gap increases. The optical band gap shift of 2.97 from 2.4 for bulk band gap is a significant shift, and is due to the quantum confinement effect observed for NCs. This shift in the band gap with change in relative ratio of polymer to NCs supports the results observed in XRD. The nanocrystal size in the present case can be calculated by the Brus equation and was found to be 6.5 nm.

Figure 8 shows the photoluminescence emission spectra of P3HT and P3HT–CdS composites in DCB solvent. A reduction in the photoluminescence intensity was observed as we increase the concentration of CdS in the solution from 1:1 to 1:8. The shape of the PL curve remains almost same; the reduction of PL intensity is enhanced with increase in CdS NC concentration in the composite matrix. This reduction of PL intensity is due to the CdS NCs formation in

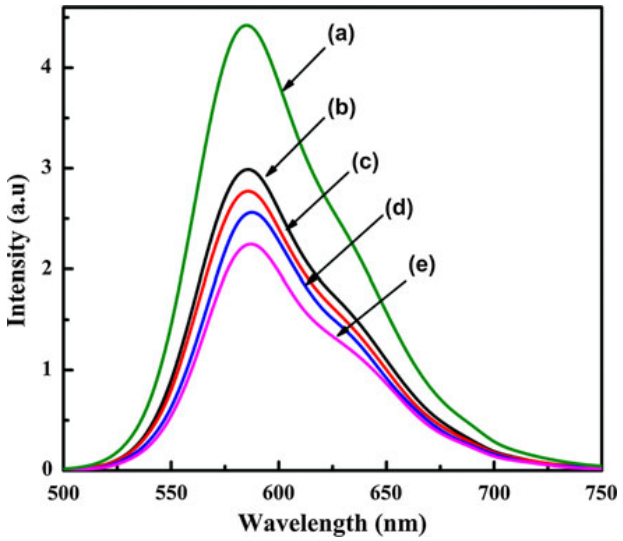


Fig. 8 PL spectra of P3HT and P3HT-CdS nanocomposites (*a*) P3HT (*b*) P3HT–CdS 1:1 (*c*) P3HT–CdS 1:2 (*d*) P3HT–CdS 1:4 (*e*) P3HT–CdS 1:8

polymer matrix and due to the photogenerated charge transfer between the CdS and P3HT. The PL quenching can be used as a powerful tool for evaluation of charge transfer from the excited polymer to the NCs, and is related to improvement in photovoltaic performance of the composite. Once the photogenerated excitons are dissociated, the probability of recombination reduces significantly. Reduced PL intensity of the composites relative to the pristine P3HT indicates that charge transfer, thereby exciton dissociation at interface of CdS and P3HT occurred. This PL quenching experiment provides us with good evidence that the polymer will be able to transfer their excited state electron to the NCs (Greenham et al. 1996). Charge transfer takes place in conjugated polymer–semiconductor NCs composites at the interface where the P3HT with a higher electron affinity transfers electron onto CdS with relatively lower electron affinity. In conversion, the polymer absorb the solar photons (charge generation), the electron is transferred to the CdS NCs and the hole potentially can remain to the polymer (charge separation). This is a well-known effect of the ultrafast electron transfer from the donor to acceptor, and it is expected to increase the exciton dissociation efficiency in PV devices. Reduction in quantum yield of polymer/nanocrystal composites compared to that of pristine P3HT, implying that a large amount of singlet excitons are not able to radiate onto ground state and they dissociate at the polymer/NCs interface. For 1:1 composite, the PL intensity decrease is about 20 % and for 1:8 composite PL decrease is about 50 %. Here, it might be observed that even after addition of a large amount of CdS NCs, the PL quenching is not complete; it might be due to some aggregation of CdS NCs in solution, as most of the NCs are not coated with polymer when CdS amount is large. PL measurements were also performed on CdS NCs synthesized in the same solvent. On excitation at 410 nm, no luminescence was detected.

PL decay

PL life time measurement provides another evidence of the charge transfer interaction in between P3HT and CdS in composite form. If the observed PL quenching for P3HT in P3HT–CdS composite is due to the charge transfer with CdS, then the interaction would accelerate the decay of the P3HT emission. To understand electron transfer dynamics, lifetime measurement studies were performed. Figure 9 shows the PL decay curves for P3HT and P3HT–CdS nanocomposites. The PL decay curves were obtained by using an excitation wavelength of 480 nm and emission was recorded at 580 nm corresponding to main P3HT peak. The rate of PL decay times at which the PL intensity decreases to $1/e$ of its initial value are used to compare the life time of P3HT and P3HT–CdS.

The emission intensity recorded at the emission maximum exhibited a mono-exponential decay

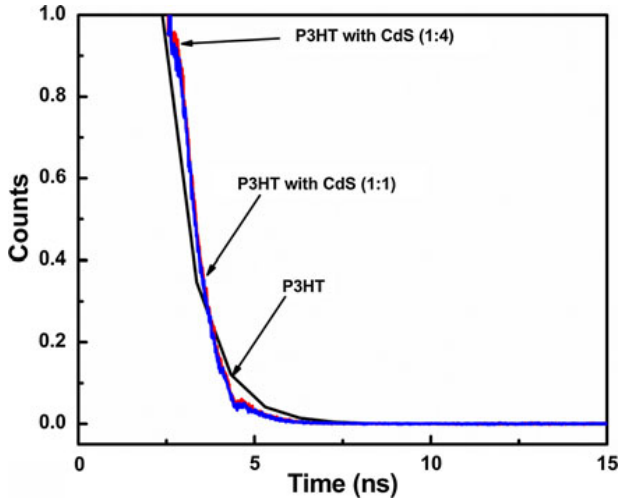


Fig. 9 Time-resolved fluorescence studies of P3HT and P3HT–CdS

Table 1 Exponential fitting results for PL decay profile of P3HT–CdS composites

Sample	A (%)	τ (ns)
P3HT	650	0.92115
P3HT–CdS (1:1)	8596	0.82137
P3HT–CdS (1:4)	9363	0.80389

$$I(t) = A\left(e^{-(t|\tau)}\right)$$

where I is the normalized emission intensity, A is pre-exponential factor, t is the time after pulsed-laser excitation, and τ is lifetime. The decays of both samples are fitted with single exponential. The kinetic parameters of the P3HT and P3HT–CdS emission decay analysis are summarized in Table 1. These results show that life time of P3HT decreased from 0.92 to 0.82 and 0.80 ns on addition of equal concentration of CdS and four times concentration of CdS with respect to P3HT concentration in P3HT–CdS composite. The P3HT–CdS composite shows a reduction of the time constant, and on increasing the amount of CdS, the time reduction is more. Further, the presence of single component in the PL decay indicates better mixing of the two components in the composite. The reduced PL decay time in the composite indicates the occurrence of a charge transfer from the excited P3HT to CdS NCs, through the transfer of electron from P3HT to CdS NCs and this indicates efficient charge transfer in P3HT–CdS nanocomposite. These results commensurate well with static PL quenching results, where a higher PL quenching was obtained for higher concentration of CdS, i.e., more quenching was observed for P3HT–CdS (1:4) composite as compared to P3HT–CdS (1:1). Therefore, it may be concluded that the charge transfer interaction actually occurred between P3HT and CdS, which provides a new path for photogenerated exciton in the composite. As a result, the chemical linkage between P3HT and CdS could facilitate the exciton dissociation for the solar cell applications.

The decrease in decay time or faster decay is due to the exciton transferring from P3HT to CdS NCs which further confirm the exciton transfer from P3HT to CdS is the main PL quenching mechanism. On increasing the concentration of CdS, not only the PL intensity decreases more, but the PL decay also becomes faster, showing more transfer of electrons due to increased CdS NCs.

Conclusions

The structural and optical properties of in situ synthesis of CdS NCs in P3HT polymer was investigated by thermolysing the single phase precursor, CdX in presence of P3HT polymer. The final size and the optical band gap of the material depend on the concentration of P3HT to CdS. The growth of CdS NCs in the presence of P3HT is hindered due to its steric effect, and arrested the growth of CdS NCs. Higher the concentration of P3HT, lower was the grain size of CdS NCs. Optical properties of the composite from the static as well as dynamic PL measurement provided evidence for charge transfer at the interface. XRD studies revealed that the CdS NCs synthesized even at such a low temperatures exist as hexagonal phase.

Acknowledgments The authors are thankful to TAPSUN (Technologies and Products for Solar Energy Utilization through Networks) program (NWP-54 project) from CSIR for financial support. The authors declare no competing financial interest.

References

Beek WJE, Wienk MM, Janssen RAJ (2006) Hybrid solar cells from regio-regular polythiophene and ZnO nanoparticles. Adv Funct Mater 16:1112–1116

Borriello C, Masala S, Bizzarro V, Nenna G, Re M, Pesce E, Minarini C, Di luccio t (2011) electroluminescence properties of poly(3-hexylthiophene)-cadmium sulfide nanoparticles grown in situ. J Appl Polym Sci 3624–3629

Coates NE, Zhou H, Kramer S, Li L, Moses D (2010) Solution based in situ synthesis and fabrication of ultrasensitive CdSe photoconductors. Adv Mater 22:5366–5369

Dayal S, Kopidakis N, Olson DC, Ginley DS, Rumbles G (2009) Direct synthesis of CdSe nanoparticles in poly(3-hexylthiophene). J Am Chem Soc 131:17726–17727

Dayal S, Kopidakis N, Olson DC, Ginley DS, Rumbles G (2010) Photovoltaic devices with a low bandgap polymer and CdSe nanostructures exceeding 3 % efficiency. Nano Lett 10:239–242

Di Luccio T, Laera AM, Tapfer L, Kempter S, Kraus R, Nickel B (2006) Controlled nucleation and growth of CdS nanoparticles in a polymer matrix. J Phys Chem B 110:12603–12609

Dowland S, Lutz T, Ward A, King SP, Sudlow A, Hill MS, Molloy KC, Haque SA (2011) Direct growth of metal sulfide nanoparticle networks in solid state polymer films for hybrid inorganic–organic solar cells. Adv Mater 23:2739–2744

Greaney MJ, Das S, Webber DH, Bradforth SE, Brutchey RL (2012) Improving open circuit potential in hybrid

P3HT:CdS bulk heterojunction solar cells via colloidal tert-butyl thiol ligand exchange. ACS Nano 6:4222–4230

Greenham NC, Peng X, Alivisatos AP (1996) Charge separation and transport in conjugated polymer/semiconductor nanocrystal composites studied by photoluminescence quenching and photoconductivity. Phys Rev B 54:17628–17637

Habas SE, Platt HAS, Maikel FAM, Van H, Ginley DS (2010) Low cost inorganic solar cells, from inks to printed device. Chem Rev 11:6571–6594

Han Z, Zhang J, Yang X, Cao W (2011) Synthesis and applications in solar cell of poly(3-octyl thiophene)/cadmium sulfide nanocomposite. Sol Energ Mat Sol C 95:483–490

Hoppe H, Sariciftci NS (2004) Organic solar cells: an overview. J Mater Res 19:1924–1945

Huynh WU, Dittmer JJ, Alivisatos AP (2002) Hybrid nanorod-polymer solar cells. Science 295:2425–2427

Jeltsch KF, Schadel M, Bonekamp JB, Niyamakom P, Rauscher F, Lademann HWA, Dumsch I, Allard S, Scherf U, Meerholz K (2012) Efficiency enhanced hybrid solar cells using a blend of quantum dots and nanorods. Adv Funct Mater 22:397–404

Khan MT, Kaur A, Dhawan SK, Chand S (2011) In situ growth of cadmium telluride nanocrystals in poly(3-hexylthiophene) matrix for photovoltaic applications. J Appl Phys 110:044509-1-044509-7

Laera AM, Resta V, Ferrara MC, Schioppa M, Piscopiello E, Tapfer L (2011) Synthesis of hybrid organic-inorganic nanocomposite materials based on CdS nanocrystals for energy conversion applications. Nanopart Res 13: 5705–5717

Leventis HC, King SP, Sudlow A, Hill MS, Molloy KC, Haque SA (2010) Nanostructured hybrid polymer-inorganic solar cell active layers formed by controllable in situ growth of semiconducting sulfide networks. Nano Lett 10:1253–1258

Li Y, Ye M, Yang C, Li X, Li Y (2005) Composition and shape controlled synthesis and optical properties of $Zn_xCd_{1-x}S$ alloyed nanocrystals. Adv Funct Mater 15:433–441

Liao HC, Chen SY, Liu DM (2009) In situ growing CdS single crystal nanorods via P3HT polymer as a soft template for enhancing photovoltaic performance. Macromolecules 42:6558–6563

Liao HC, Lin CC, Chen YW, Liu TC, Chen SY (2010) Improvement in photovoltaic performance for hybrid P3HT/elongated CdS nanocrystals solar cells with F-doped SnO_2 arrays. J Mater Chem 20:5429–5435

Liao HC, Chantarat N, Chen SY, Peng CH (2011) Annealing effect on photovoltaic performance of hybrid P3HT/in-situ grown CdS nanocrystal solar cell. J Electrochem Soc 158(7):E67–E72

Liu W (2006) Low temperature synthesis of hexagonal phase ZnS nanocrystals by thermolysis of an air stable single source molecular precursor in air. Mater Lett 60:551–554

Liu J, Wang W, Yu H, Wu Z, Peng J, Cao Y (2008) Surface ligand effect in MEHPPV/TiO_2 hybrid solar cells. Sol Energy Mater Sol C 92:1403–1409

Maier E, Rath T, Hass W, Werzer O, Saf R, Hofer F, Meissner D, Volobujeva O, Bereznev S, Mellikov E, Amenitsch H, Resel R, Trimmel G (2011) $CuInS_2$-poly(3-ethyl-4-butanoate)thiophene nanocomposite solar cells: preparation by an in situ formation route, performance and stability issues. Sol Energy Mater Sol C 95:1354–1361

Masala S, Gobbo SD, Borriello C, Bizzarro V, Ferrara VL, Pesce E, Minarini C, Crescenzi MD, Di Luccio T (2011) Hybrid polymer–CdS solar cell active layers formed by in situ growth of CdS nanoparticles. J Nanopart Res 13:6537–6544

Mayukh M, Jung IH, He F, Yu L (2012) Incremental optimization in donor polymers for bulk heterojunction organic solar cells exhibiting high performance. J Polym Sci 50:1057–1070

Moule AJ, Chang L, Thambidurai C, Vidu R, Stroeve P (2012) Hybrid solar cells, basic principles and role of ligands. J Mater Chem 22:2351–2368

Nair PS, Radhakrishnan T, Revaprasadu N, Kolawole GA, Brien PO (2002) A novel single source precursor for the preparation of CdS nanoparticles. J Mater Chem 12:2722–2725

Nair PS, Radhakrishna T, Revaprasadu N, Kolawole GA, Luyt AS, Djokovic VJ (2005) Preparation and characterization of polystyrene films containing PbS nanoparticles. Mater Sci 40:4407–4409

Plante IJL, Zeid TW, Yang P, Mokari T (2010) Synthesis of metal sulfide nanomaterials via thermal decomposition of single source precursors. J Mater Chem 20:6612–6617

Pradhan N, Katz B, Efrima S (2003) Synthesis of high-quality metal sulfide nanoparticles from alkyl xanthate single precursors in alkylamine solvents. J Phys Chem B107:13843–13854

Reiss P, Couderc E, Girolamo JD, Pron A (2010) Confined polymers/semiconductor nanocrystals hybrid materials-preparation, electrical transport properties and applications. Nanoscale 3:446–489

Ren S, Chang LY, Lim SK, Zhao J, Smith M, Ni Z, Bulovic V, Bawendi M, Gradecak S (2011) Inorganic–organic hybrid solar cell: bridging quantum dots to conjugated polymer nanowires. Nano Lett 11:3998–4002

Resta V, Laera AM, Piscopiello E, Schioppa M, Tapfer L (2010) Highly efficient precursors for direct synthesis of tailored CdS nanocrystals in organic polymers. J Phys Chem C 114:17311–17317

Saunders BR, Turner ML (2008) Nanoparticle–polymer photovoltaic cell. Adv Colloid Interface Sci 138:1–23

Shen S, Zhang Y, Peng L, Xu B, Du Y, Deng M, Xu H, Wang Q (2011) Generalized synthesis of metal sulfide nanocrystals from single source precursors: size, shape and chemical composition control and their properties. Crys Eng Commun 13:4572–4579

Stavrinadis A, Beal R, Smith JM, Assender HE, Watt AAR (2008) Direct formation of PbS nanorods in a conjugated polymer. Adv Mater 20:3105–3109

Talapin DV, Lee JS, Kovalenko MV, Shevchenko EV (2010) Prospects of colloidal nanocrystals for electronic and optoelectronic applications. Chem Rev 110:389–458

Verma D, Rao AR, Dutta V (2009) Surfactant free CdTe nanoparticles mixed MEHPP hybrid solar cell deposited by spin coating technique. Sol Energy Mater Sol C 93(9):1482–1487

Wang Z, Qu S, Zeng X, Zhang C, Shi M, Tan F, Wang Z, Liu J, Hou Y, Teng F, Feng Z (2008) Synthesis of MDMO-PPV capped PbS quantum dots and their application to solar cells, polymer. Polymer 49(21):4647–4651

Wang Y, Wei W, Liu X, Gu Y (2012) Research progress on polymer heterojunction solar cells, solar energy materials and solar cells. Sol Energy Mater Sol C 98:129–145

Warner JH, Watt AAA (2006) Monodisperse PbS nanocrystals synthesized in a conducting polymer. Mater Lett 60:2375–2379

Zhou Y, Eck M, Kruger M (2010) Bulk-heterojunction hybrid solar cells based on colloidal nanocrystals and conjugated polymers. Energy Environ Sci 3:1851–1864